PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47, A61K 38/17, C07K 16/18, C12Q 1/68

A2

(11) International Publication Number: WO 99/45114

(43) International Publication Date: 10 September 1999 (10.09.99)

(21) International Application Number: PCT/US99/04758

(22) International Filing Date: 3 March 1999 (03.03.99)

(30) Priority Data:
60/076,611 3 March 1998 (03.03.98) US

(71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(72) Inventors: HOLLOWAY, James, L.; 835 N.E. 89th Street, Seattle, WA 98115 (US). LOFTON-DAY, Catherine, E.; 23908 35th Avenue West, Brier, WA 98036 (US).

(74) Agent: LINGENFELTER, Susan, E.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN SEMAPHORIN ZSMF-7

(57) Abstract

Semaphorin polypeptides, polynucleotides encoding the polypeptides, and related compositions and methods are disclosed. The polypeptides are expressed in neuronal and lymphatic tissues. The polypeptides may be used within methods for detecting receptors that mediate neurite outgrowth, modulate cellular proliferation and/or differentiation, and immune response.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

DESCRIPTION

HUMAN SEMAPHORIN ZSMF-7

BACKGROUND OF THE INVENTION

Neuronal cell outgrowths, known as processes, grow away from the cell body to form synaptic connections. Long, thin processes which carry information away from the cell body are called axons, and short, thicker processes which carry information to and from the cell body are called dendrites. Axons and dendrites are collectively referred to as neurites. Neurites are extended by means of growth cones, the growing tip of the neurite, which is highly motile and is ultimately responsible for increasing and extending the neuronal network in the body. The growth cones are able to navigate their way to their targets using environmental cues or signals, which encourage or discourage the growth cone from extending the neurite in a particular direction. Such cues and signals include older neurons and orienting glial fibers, chemicals such as nerve growth factor released by astrocytes and other attracting or repelling substances released by target cells. The membrane of the growth cone bears molecules such as N-CAM (nerve cell adhesion molecule) which are attracted or repelled by environmental cues and thus influence the direction and degree of neurite growth. The growth cone also engulfs molecules from the environment which are transported to the cell body and influence growth. A number of proteins from vertebrates and invertebrates have been identified as influencing the guidance of neurite growth, either through repulsion or chemoattraction. Among those molecules are netrins, EPH-related receptor tyrosine kinases and their ligands, vitronectin, thrombospondin, human neuronal attachment factor-1 (NAF-1), connectin, adhesion molecules such as CAM (cell adhesion molecule) and

the semaphorins/collapsins (Neugebauer et al., Neuron 6:345-58, 1991; O'Shea et al., Neuron 7:231-7, 1991; Osterhout et al., Devel. Biol. 150:256-65, 1992; Goodman, Cell 78:353-6, 1993; DeFreitas et al., Neuron 15:333-43, 1995; Dodd and Schuchardy Cell 81:471-4, 1995; Keynes and Cook, Cell 83:161-9, 1995; Müller et al., Cur. Opin. Genet. and Devel. 6:469-74, 1996, Goodman, Annu. Rev. Neurosci. 19:341-77, 1996; WIPO Patent Application No: 97/29189 and Goodman et al., US Patent No. 5,639,856).

Semaphorins/collapsins are a family of related transmembrane and secreted molecules. Invertebrate, vertebrate and viral semaphorins are known (Kolodkin et al., Cell 75:1389-99, 1993; Luo et al., Cell 75:217-27, 1993; Ensser and Fleckenstein, J. Gen. Virol. 76:1063-7, 1995; Luo et al, Neuron 14:1131-40, 1995; Adams et al., Mech. Devel. 57:33-45, 1996; Hall et al., Proc. Natl. Acad. Sci. USA 93:11780-8, 1996; Roche et al., Oncogene 12:1289-97, 1996; Skeido et al., Proc. Natl. Acad. Sci. USA 93:4120-5, 1996; Xiang et al., Genomics 32:39-48, 1996; Eckhardt et al., Mol. Cell Neurosci. 9:409-19, 1997 and Zhou et al., Mol. Cell. Neurosci. 9:26-41, 1997).

The semaphorins generally comprise an N-terminal variable region of 30-60 amino acids that includes a secretory signal sequence, followed by a conserved region of about 500 amino acid residues called the semaphorin or sema domain. The extracellular semaphorin domain contains between 13-16 conserved cysteine residues, an N-linked glycosylation site and numerous blocks of amino acid residues which are conserved though-out the family. Classification into five subgroups within the semaphorin family has made based on the sequence of the region C-terminal to the semaphorin domain. Both soluble (lacking a transmembrane domain) and membrane-bound (having a transmembrane domain and localized to a membrane) semaphorins have been described. See, for example, Kolodkin et al., ibid.; Adams et al., ibid. and Goodman et al., US Patent No:5,639,856.

Group I semaphorins include semaphorins having a transmembrane domain followed by a cytoplasmic domain. Most insect semaphorins are membrane bound proteins and belong to Group I. G-Sema I, T-Sema I and D-Sema I have a region of 80 amino acid residues following the semaphorin domain, which is followed by a transmembrane domain and an 80-110 amino acid cytoplasmic domain. Murine Sema IVa has a transmembrane domain followed by a 216 amino acid cytoplasmic domain.

Groups II and III have no transmembrane domain or membrane association, but have a region with Ig homology. Group II secreted proteins, such as D-sema II, have a region of less than 20 amino acids between the semaphorin domain and an Ig-like domain followed by a short region of amino acid residues. Also included is alcelaphine herpesvirus type 1 semaphorin-like gene (avh-sema, Ensser and Fleckenstein, J. Gen. Virol. 76:1063-7, 1995) which ends with an Ig-like domain. Group III proteins, such as H-Sema III, are similar to Group II with the exception that the C-terminal amino acid region following the Ig-like domain is longer.

Group IV has a region of Ig homology C-terminal of the semaphorin domain followed by a transmembrane and cytoplasmic domain and includes semaphorins such as Sem B.

Group V has a series of thrombospondin repeats C-terminal of the semaphorin domain followed by a transmembrane and cytoplasmic domain and include murine sema F and G.

Other viral semaphorins such as vaccinia virus sema IV and variola virus sema IV, have a truncated, 441 amino acid residue, semaphorin domain and no Ig region. See Kolodkin et al., ibid.; Adams et al. ibid. and Zhou et al. ibid.

Overall semaphorins share the greatest degree of homology within the semaphorin domain, between, 25-93%, and a greater degree of divergence in all other regions and domains, suggesting distinct roles for various sub-groups

within the semaphorin family. The viral semaphorins are the most diverse, sharing only 25% identity with vertebrate semaphorins. Between vertebrate and invertebrate semaphorins, the percent identity varies between 30-40%.

Neurite growth cues are of great therapeutic value. Isolating and characterizing novel semaphorins would be of value for example, in modulating neurite growth and development; treatment of peripheral neuropathies; for use as therapeutics for the regeneration of neurons following strokes, brain damage caused by head injuries and paralysis caused by spinal injuries; diagnosing neurological diseases and in treating neurodegenerative diseases such as multiple sclerosis, Alzheimer's disease and Parkinson's disease. In addition, semaphorins are also being found in non-neuronal tissues and their usefulness for modulating cellular proliferation and differentiation as well as mediating immunological responses is now being reported. The present invention addresses these needs and others by providing novel semaphorins and related compositions and methods.

SUMMARY OF THE INVENTION

The present invention provides a novel semaphorin polypeptide and related compositions and methods.

Within one aspect is provided an isolated semaphorin polypeptide comprising a sequence of amino acid residues that is at least 80% identical in amino acid sequence to residues 69 through 541 of SEQ ID NO:2, said polypeptide comprising cysteine residues at positions corresponding to residues 126, 143, 152, 266, 291, 335, 366, 493, 500, 503, 511, 518 and 541 of SEQ ID NO:2. Within one embodiment the sequence of amino acid residues is at least 90% identical. Within another embodiment the polypeptide further comprises an Ig-like domain. Within a related embodiment the Ig-like domain comprises a sequence of amino acids from residue 561-620 of SEQ ID NO:2. Within another embodiment the polypeptide comprises residues 45-666 of SEQ ID NO:2. Within yet another embodiment the sequence of amino acid residues comprises residues 1-666 of SEQ ID NO:2. Within another embodiment the sequence of amino acid residues is from 473-624 amino acid residues. The invention further provides an isolated semaphorin polypeptide selected from the group consisting of: a) a polypeptide comprising a sequence of amino acid residues from amino acid residue 45 to residue 666 of SEQ ID NO:2; b) a polypeptide comprising a sequence of amino acid residues from amino acid residue 69 to residue 666 of SEQ ID NO:2; c) a polypeptide comprising a sequence of amino acid residues from amino acid residue 69 to residue 541 of SEQ ID NO:2 and d) a polypeptide comprising a sequence of amino acid residues from amino acid residue 1 to residue 666 of SEQ ID NO:2. Within yet another embodiment any difference between said amino acid sequence of said isolated polypeptide and said corresponding amino acid sequence of SEQ ID NO:2 is due to a conservative amino acid substitution. Within another embodiment the polypeptide is covalently linked to a moiety selected from the group consisting of affinity tags, toxins, radionucleotides,

enzymes and fluorophores. Within a related embodiment the moiety is an affinity tag selected from the group consisting of polyhistidine, FLAG, Glu-Glu, glutathione S transferase and an immunoglobulin heavy chain constant region. Within a further related embodiment the polypeptide further comprises a proteolytic cleavage site between said sequence of amino acid residues and said affinity tag.

Within another aspect the invention provides an expression vector comprising the following operably linked elements: a transcription promoter; a DNA segment encoding a semaphorin polypeptide comprising a sequence of amino acid residues that is at least 80% identical in amino acid sequence to residues 69 through 541 of SEQ ID NO:2, said polypeptide comprising cysteine residues at positions corresponding to residues 126, 143, 152, 266, 291, 335, 366, 493, 500, 503, 511, 518 and 541 of SEQ ID NO:2; and a transcriptional terminator. Within one embodiment the expression vector further comprises a secretory signal sequence operably linked to said DNA segment. Within a related embodiment the secretory signal sequence encodes residues 1-44 of SEQ ID NO:2. Within another embodiment the sequence of amino acid residues is at least 90% identical. Within another embodiment the DNA segment encodes a semaphorin polypeptide comprising an Ig-like domain. Within a related embodiment the Ig-like domain comprises a sequence of amino acids from residue 561-620 of SEQ ID NO:2. Within another embodiment the sequence of amino acid residues comprises residues 45-666 of SEQ ID NO:2. Within yet another embodiment the DNA segment encodes a semaphorin polypeptide covalently linked to an affinity tag selected from the group consisting of polyhistidine, FLAG, Glu-Glu, glutathione S transferase and an immunoglobulin heavy chain constant region. The invention further provides a cultured cell into which has been introduced an expression vector as described above, wherein said cell expresses the polypeptide encoded by the

DNA segment. The invention also provides a method of producing a semaphorin protein comprising: culturing a cell into which has been introduced an expression vector as described above, whereby said cell expresses said semaphorin protein encoded by said DNA segment; and recovering said expressed semaphorin protein.

Within another aspect the invention provides a pharmaceutical composition comprising a polypeptide as described above, in combination with a pharmaceutically acceptable vehicle.

Within another aspect the invention provides an antibody or antibody fragment that specifically binds to an epitope of a semaphorin polypeptide as described above. Within one embodiment the antibody is selected from the group consisting of: a) polyclonal antibody; b) murine monoclonal antibody; c) humanized antibody derived from b); and d) human monoclonal antibody. Within a related embodiment the antibody fragment is selected from the group consisting of F(ab'), F(ab), Fab', Fab, Fv, scFv, and minimal recognition unit. Within a related embodiment is provided an anti-idiotype antibody that specifically binds to the antibody described above.

Within another aspect the invention provides a binding protein that specifically binds to an epitope of a semaphorin polypeptide as described above.

Within yet another aspect the invention provides an isolated polynucleotide encoding a semaphorin polypeptide comprising a sequence of amino acid residues that is at least 80% identical in amino acid sequence to residues 69 through 541 of SEQ ID NO:2, said polypeptide comprising cysteine residues at positions corresponding to residues 126, 143, 152, 266, 291, 335, 366, 493, 500, 503, 511, 518 and 541 of SEQ ID NO:2. Within one embodiment the sequence of amino acid residues is at least 90% identical. Within another embodiment the semaphorin polypeptide comprises an Ig-like domain. Within a related embodiment the Ig-like domain comprises a sequence of amino acids from

residue 561-620 of SEQ ID NO:2. Within another embodiment the sequence of amino acid residues comprises residues 45-666 of SEQ ID NO:2. Within yet another embodiment the sequence of amino acid residues comprises residues 1-666 of SEQ ID NO:2. Within another embodiment the polynucleotide comprises nucleotide 1 to nucleotide 1998 of SEQ ID NO:5. Also provided by the invention is an isolated polynucleotide selected from the group consisting of: a)

a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 152 to nucleotide 2017 of SEQ ID NO:1; b) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 244 to nucleotide 2017 of SEQ ID NO:1; c) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 244 to nucleotide 1640 of SEQ ID NO:1; d) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 20 to nucleotide 2017 of SEQ ID NO:1; and e) a complementary polynucleotide sequence of a, b, c or d.

Within another aspect the invention provides a method for detecting a genetic abnormality in a patient, comprising: obtaining a genetic sample from a patient; incubating the genetic sample with a polynucleotide comprising at least 14 contiguous nucleotides of SEQ ID NO:1 or the complement of SEQ ID NO:1, under conditions wherein said polynucleotide will hybridize to complementary polynucleotide sequence, to produce a first reaction product; comparing said first reaction product to a control reaction product, wherein a difference between said first reaction product and said control reaction product is indicative of a genetic abnormality in the patient.

These and other aspects of the invention will become evident upon reference to the following detailed description and the attached drawing.

BRIEF DESCRIPTION OF THE DRAWING

The figure shows an alignment of ZSMF-7 (SEQ ID NO:2), alcelaphine herpesvirus type 1 semaphorin-like gene (AHU18243) (SEQ ID NO:31), mouse semaA (SEQ ID NO:33), mouse semaB (SEQ ID NO:3), mouse semaC (SEQ ID NO:30), mouse semaD (SEQ ID NO:32), mouse semaE (SEQ ID NO:29) and mouse semaF (SEQ ID NO:23) is shown in the Figure. There are clusters of conserved or highly homologous amino acids throughout the semaphorin domains of these semaphorin proteins. Conserved amino acid residues are indicated by "*" and residues with a high degree of homology are indicated by ":" and ".".

DETAILED DESCRIPTION OF THE INVENTION

Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms to be used hereinafter:

The term "affinity tag" is used herein to denote a polypeptide segment that can be attached to a second polypeptide to provide for purification of the second polypeptide or provide sites for attachment of the second polypeptide to a substrate. In principal, any peptide or protein for which an antibody or other specific binding agent is available can be used as an affinity tag. Affinity tags include a poly-histidine tract, protein A (Nilsson et al., EMBO J. 4:1075, 1985; Nilsson et al., Methods Enzymol. 198:3, 1991), glutathione S transferase (Smith and Johnson, Gene 67:31, 1988), Glu-Glu affinity tag (Grussenmeyer et al., Proc. Natl. Acad. Sci. USA 82:7952-4, 1995), substance P, Flag™ peptide (Hopp et al., Biotechnology 6:1204-10, 1988), streptavidin binding peptide, or other antigenic epitope or binding domain. See, in general, Ford et al., Protein Expression and Purification 2: 95-107, 1991. DNAs encoding affinity tags are available from commercial suppliers (e.g., Pharmacia Biotech, Piscataway, NJ).

The term "allelic variant" is used herein to denote any of two or more alternative forms of a gene

occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene.

The terms "amino-terminal" and "carboxyl-terminal" are used herein to denote positions within polypeptides. Where the context allows, these terms are used with reference to a particular sequence or portion of a polypeptide to denote proximity or relative position. For example, a certain sequence positioned carboxyl-terminal to a reference sequence within a polypeptide is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete polypeptide.

The term "complements of a polynucleotide molecule" is a polynucleotide molecule having a complementary base sequence and reverse orientation as compared to a reference sequence. For example, the sequence 5' ATGCACGGG 3' is complementary to 5' CCCGTGCAT 3'.

The term "contig" denotes a polynucleotide that has a contiguous stretch of identical or complementary sequence to another polynucleotide. Contiguous sequences are said to "overlap" a given stretch of polynucleotide sequence either in their entirety or along a partial stretch of the polynucleotide. For example, representative contigs to the polynucleotide sequence 5'-ATGGCTTAGCTT-3' are 5'-TAGCTTgagtct-3' and 3'-gtcgacTACCGA-5'.

The term "degenerate nucleotide sequence" denotes a sequence of nucleotides that includes one or more degenerate codons (as compared to a reference polynucleotide molecule that encodes a polypeptide). Degenerate codons contain different triplets of

nucleotides, but encode the same amino acid residue (i.e., GAU and GAC triplets each encode Asp).

The term "expression vector" is used to denote a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that provide for its transcription. Such additional segments include promoter and terminator sequences, and may also include one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, etc. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both.

The term "isolated", when applied to a polynucleotide, denotes that the polynucleotide has been removed from its natural genetic milieu and is thus free of other extraneous or unwanted coding sequences, and is in a form suitable for use within genetically engineered protein production systems. Such isolated molecules are those that are separated from their natural environment and include cDNA and genomic clones. Isolated DNA molecules of the present invention are free of other genes with which they are ordinarily associated, but may include naturally occurring 5' and 3' untranslated regions such as promoters and terminators. The identification of associated regions will be evident to one of ordinary skill in the art (see for example, Dynan and Tijan, <u>Nature</u> <u>316</u>:774-78, 1985).

An "isolated" polypeptide or protein is a polypeptide or protein that is found in a condition other than its native environment, such as apart from blood and animal tissue. In a preferred form, the isolated polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin. It is preferred to provide the polypeptides in a highly purified form, i.e. greater than 95% pure, more preferably greater than 99% pure. When used in this context, the term "isolated" does not exclude the presence of the same

polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

The term "operably linked", when referring to DNA segments, indicates that the segments are arranged so that they function in concert for their intended purposes, e.g., transcription initiates in the promoter and proceeds through the coding segment to the terminator.

The term "ortholog" denotes a polypeptide or protein obtained from one species that is the functional counterpart of a polypeptide or protein from a different species. Sequence differences among orthologs are the result of speciation.

A "polynucleotide" is a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized *in vitro*, or prepared from a combination of natural and synthetic molecules. Sizes of polynucleotides are expressed as base pairs (abbreviated "bp"), nucleotides ("nt"), or kilobases ("kb"). Where the context allows, the latter two terms may describe polynucleotides that are single-stranded or double-stranded. When the term is applied to double-stranded molecules it is used to denote overall length and will be understood to be equivalent to the term "base pairs". It will be recognized by those skilled in the art that the two strands of a double-stranded polynucleotide may differ slightly in length and that the ends thereof may be staggered as a result of enzymatic cleavage; thus all nucleotides within a double-stranded polynucleotide molecule may not be paired. Such unpaired ends will in general not exceed 20 nt in length.

A "polypeptide" is a polymer of amino acid residues joined by peptide bonds, whether produced naturally or synthetically. Polypeptides of less than about 10 amino acid residues are commonly referred to as "peptides".

"Probes and/or primers" as used herein can be RNA or DNA. DNA can be either cDNA or genomic DNA. Polynucleotide probes and primers are single or double-stranded DNA or RNA, generally synthetic oligonucleotides, but may be generated from cloned cDNA or genomic sequences or its complements. Analytical probes will generally be at least 20 nucleotides in length, although somewhat shorter probes (14-17 nucleotides) can be used. PCR primers are at least 5 nucleotides in length, preferably 15 or more nt, more preferably 20-30 nt. Short polynucleotides can be used when a small region of the gene is targeted for analysis. For gross analysis of genes, a polynucleotide probe may comprise an entire exon or more. Probes can be labeled to provide a detectable signal, such as with an enzyme, biotin, a radionuclide, fluorophore, chemiluminescer, paramagnetic particle and the like, which are commercially available from many sources, such as Molecular Probes, Inc., Eugene, OR, and Amersham Corp., Arlington Heights, IL, using techniques that are well known in the art. Examples of ZSMF-7 probes and primers include, but are not limited to, the sequences disclosed herein as SEQ ID NOs: 4, 6, 7, 9-21, 24, 25, 26 and 28.

The term "promoter" is used herein for its art-recognized meaning to denote a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

A "protein" is a macromolecule comprising one or more polypeptide chains. A protein may also comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a protein by the cell in which the protein is produced, and will vary with the type of cell. Proteins are defined herein in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless.

The term "receptor" denotes a cell-associated protein that binds to a bioactive molecule (i.e., a ligand) and mediates the effect of the ligand on the cell. Membrane-bound receptors are characterized by a multi-domain structure comprising an extracellular ligand-binding domain and an intracellular effector domain that is typically involved in signal transduction. Binding of ligand to receptor results in a conformational change in the receptor that causes an interaction between the effector domain and other molecule(s) in the cell. This interaction in turn leads to an alteration in the metabolism of the cell. Metabolic events that are linked to receptor-ligand interactions include gene transcription, phosphorylation, dephosphorylation, increases in cyclic AMP production, mobilization of cellular calcium, mobilization of membrane lipids, cell adhesion, hydrolysis of inositol lipids and hydrolysis of phospholipids. In general, receptors can be membrane bound, cytosolic or nuclear; monomeric (e.g., thyroid stimulating hormone receptor, beta-adrenergic receptor) or multimeric (e.g., PDGF receptor, growth hormone receptor, IL-3 receptor, GM-CSF receptor, G-CSF receptor, erythropoietin receptor and IL-6 receptor).

The term "secretory signal sequence" denotes a DNA sequence that encodes a polypeptide (a "secretory peptide") that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

The term "splice variant" is used herein to denote alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode

PAGES

15 – 50

NOT TO BE TAKEN INTO ACCOUNT FOR
THE PURPOSES OF INTERNATIONAL PROCESSING

directing and defining the growth of developing tissue, in particular, defining the margins of a particular organ or tissue. ZSMF-7 polypeptides would be useful in the defining and directing development of various tissues and organs including those associated with muscle, fibroblasts, reproductive, endocrine and lymphatic.

Semaphorins have also been associated with non-neuronal functions. Viral semaphorins have been speculated to act as modulators of the immune system, as natural immunosuppressants reducing the immune response by mimicking the function of a particular subfamily of semaphorins that can modulate immune functions (Kolodkin et al., ibid., and Ensser and Fleckenstein, ibid.). Other non-viral semaphorins are also associated with the immune system. Human semaphorin E, which is homologous to viral cytokine inhibiting proteins, contains conserved regions of amino acid residues that have been found in the viral semaphorins. Semaphorin E was found to be upregulated in rheumatoid synovial fibroblastoid cells which suggests that it may have a role as a regulator of inflammatory processes and an involvement in the development of rheumatoid arthritis (Mangasser-Stephan et al., Biochem. Biophys. Res. Comm. 234:153-6, 1997). Semaphorin CD100 has been reported to promote B-cell growth and aggregation and may be involved in lymphocyte activation (Hall et al., Proc. Natl. Acad. Sci. USA 93: 11780-5, 1996) and its mouse homologue, mSema G, is expressed on lymphocytes and is suggested to play a role in the immune system as well (Furuyama et al., J. Biol. Chem. 271:33376-81, 1996).

ZSMF-7 shares the greatest homology with a viral semaphorin, alcelaphine herpesvirus type 1 semaphorin-like gene (ahv-sema) and coupled with the strong mRNA expression in activated T lymphocytes suggests that ZSMF-7 plays a role as a mediator of immunosuppression, in particular the activation and regulation of T lymphocytes. ZSMF-7 polypeptides would be useful additions to therapies for treating immunodeficiencies. ZSMF-7 was expressed in

activated lymphocytes (MRL cells) and not in resting lymphocyte cells ($CD4^+$ and $CD8^+$) suggesting that it would be useful tool for diagnosis and treatment of conditions where selective elimination of inappropriately activated T cells would be benificial, such as in autoimmune diseases, in particular insulin dependent diabetes mellitus, rheumatoid arthritis and multiple sclerosis. Such polypeptides could be used to screen serum samples from patients suffering from such conditions. Inappropriately activated T cells would include those specific for self-peptide/self-major histocompatibility complexes and those specific for non-self antigens from transplanted tissues. Use could also be made of these polypeptides in blood screening for removal of inappropriately activated T cells before returning the blood to the donor. Those skilled in the art will recognize that conditions related to ZSMF-7 underexpression or overexpression may be amenable to treatment by therapeutic manipulation of ZSMF-7 protein levels.

ZSMF-7 polypeptides can be used *in vivo* as an anti-inflammatory, for inhibition of antigen in humoral and cellular immunity and for immunosuppression in graft and organ transplants.

ZSMF-7 polynucleotides and/or polypeptides can be used for regulating the proliferation and stimulation of a wide variety of cells, such as T cells, B cells, lymphocytes, peripheral blood mononuclear cells, fibroblasts and hematopoietic cells. ZSMF-7 polypeptides will also find use in mediating metabolic or physiological processes *in vivo*. Proliferation and differentiation can be measured *in vitro* using cultured cells. Suitable cell lines are available commercially from such sources as the American Type Culture Collection (Rockville, MD). Bioassays and ELISAs are available to measure cellular response to ZSMF-7, in particular are those which measure changes in cytokine production as a measure of cellular response (see for example, Current Protocols in Immunology ed. John Coligan et al., NIH, 1996). Also of interest are

apoptosis assays, such as the DNA fragmentation assay described by Wiley et al. (Immunity, 3:673-82, 1995, and the cell death assay described by Pan et al., Science, 276:111-13, 1997). Assays to measure other cellular responses, including antibody isotype, monocyte activation, NK cell formation and antigen presenting cell function are also known. The ZSMF-7 polypeptides may also be used to stimulate lymphocyte development, such as during bone marrow transplantation and as therapy for some cancers.

In vivo response to ZSMF-7 polypeptides can also be measured by administering polypeptides of the claimed invention to the appropriate animal model. Well established animal models are available to test *in vivo* efficacy of ZSMF-7 polypeptides for certain disease states. In particular, ZSMF-7 polypeptides can be tested *in vivo* in a number of animal models of autoimmune disease, such as the NOD mice, a spontaneous model system for insulin-dependent diabetes mellitus (IDDM), to study induction of non-responsiveness in the animal model. Administration of ZSMF-7 polypeptides prior to or after onset of disease can be monitored by assay of urine glucose levels in the NOD mouse. Alternatively, induced models of autoimmune disease, such as experimental allergic encephalitis (EAE), can be administered ZSMF-7 polypeptides. Administration in a preventive or intervention mode can be followed by monitoring the clinical symptoms of EAE. In addition, ZSMF-7 polypeptides can be tested *in vivo* in animal models for cancer, where suppression or apoptosis of introduced tumor cells can be monitored following administration of ZSMF-7.

The present invention also provides reagents for use in diagnostic applications. For example, the ZSMF-7 gene, a probe comprising ZSMF-7 DNA or RNA, or a subsequence thereof can be used to determine if the ZSMF-7 gene is present on chromosome 15 or if a mutation has occurred. Detectable chromosomal aberrations at the ZSMF-7 gene locus include, but are not limited to, aneuploidy,

gene copy number changes, insertions, deletions, restriction site changes and rearrangements. These aberrations can occur within the coding sequence, within introns, or within flanking sequences, including upstream promoter and regulatory regions, and may be manifested as physical alterations within a coding sequence or changes in gene expression level. Deletion of the region 3p21, associated with human semaphorin III/F (also known as human semaphorin IV), is correlated with small cell lung cancer (Roche et al., Oncogene 12:1289-97, 1996 and Xiang et al., Genomics 32:39-48, 1996).

In general, these diagnostic methods comprise the steps of (a) obtaining a genetic sample from a patient; (b) incubating the genetic sample with a polynucleotide probe or primer as disclosed above, under conditions wherein the polynucleotide will hybridize to complementary polynucleotide sequence, to produce a first reaction product; and (iii) comparing the first reaction product to a control reaction product. A difference between the first reaction product and the control reaction product is indicative of a genetic abnormality in the patient. Genetic samples for use within the present invention include genomic DNA, cDNA, and RNA. The polynucleotide probe or primer can be RNA or DNA, and will comprise a portion of SEQ ID NO:1, the complement of SEQ ID NO:1, or an RNA equivalent thereof. Suitable assay methods in this regard include molecular genetic techniques known to those in the art, such as restriction fragment length polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, ligation chain reaction (Barany, PCR Methods and Applications 1:5-16, 1991), ribonuclease protection assays, and other genetic linkage analysis techniques known in the art (Sambrook et al., ibid.; Ausubel et. al., ibid.; Marian, Chest 108:255-65, 1995). Ribonuclease protection assays (see, e.g., Ausubel et al., ibid., ch. 4) comprise the hybridization of an RNA probe to a patient RNA sample, after which the reaction

product (RNA-RNA hybrid) is exposed to RNase. Hybridized regions of the RNA are protected from digestion. Within PCR assays, a patient's genetic sample is incubated with a pair of polynucleotide primers, and the region between the primers is amplified and recovered. Changes in size or amount of recovered product are indicative of mutations in the patient. Another PCR-based technique that can be employed is single strand conformational polymorphism (SSCP) analysis (Hayashi, PCR Methods and Applications 1:34-8, 1991).

As a ligand, the activity of ZSMF-7 polypeptide can be measured by a silicon-based biosensor microphysiometer which measures the extracellular acidification rate or proton excretion associated with receptor binding and subsequent physiologic cellular responses. An exemplary device is the Cytosensor™ Microphysiometer (Molecular Devices, Sunnyvale, CA). A variety of cellular responses, such as cell proliferation, ion transport, energy production, inflammatory response, regulatory and receptor activation, and the like, can be measured by this method. See, for example, McConnell et al., Science 257:1906-12, 1992; Pitchford et al., Meth. Enzymol. 228:84-108, 1997; Arimilli et al., J. Immunol. Meth. 212:49-59, 1998; Van Liefde et al., Eur. J. Pharmacol. 346:87-95, 1998. The microphysiometer can be used for assaying adherent or non-adherent eukaryotic or prokaryotic cells. By measuring extracellular acidification changes in cell media over time, the microphysiometer directly measures cellular responses to various stimuli, including ZSMF-7 polypeptide, its agonists, or antagonists. Preferably, the microphysiometer is used to measure responses of a ZSMF-7-responsive eukaryotic cell, compared to a control eukaryotic cell that does not respond to ZSMF-7 polypeptide. ZSMF-7-responsive eukaryotic cells comprise cells into which a receptor for

ZSMF-7 has been transfected creating a cell that is responsive to ZSMF-7; or cells naturally responsive to ZSMF-7 such as cells derived from neurological, endrocrinological or tumor tissue. Differences, measured by a change, for example, an increase or diminution in extracellular acidification, in the response of cells exposed to ZSMF-7 polypeptide, relative to a control not exposed to ZSMF-7, are a direct measurement of ZSMF-7-modulated cellular responses. Moreover, such ZSMF-7-modulated responses can be assayed under a variety of stimuli. Using the microphysiometer, there is provided a method of identifying agonists of ZSMF-7 polypeptide, comprising providing cells responsive to a ZSMF-7 polypeptide, culturing a first portion of the cells in the absence of a test compound, culturing a second portion of the cells in the presence of a test compound, and detecting a change, for example, an increase or diminution, in a cellular response of the second portion of the cells as compared to the first portion of the cells. The change in cellular response is shown as a measurable change extracellular acidification rate. Moreover, culturing a third portion of the cells in the presence of ZSMF-7 polypeptide and the absence of a test compound can be used as a positive control for the ZSMF-7-responsive cells, and as a control to compare the agonist activity of a test compound with that of the ZSMF-7 polypeptide. Moreover, using the microphysiometer, there is provided a method of identifying antagonists of ZSMF-7 polypeptide, comprising providing cells responsive to a ZSMF-7 polypeptide, culturing a first portion of the cells in the presence of ZSMF-7 and the absence of a test compound, culturing a second portion of the cells in the presence of ZSMF-7 and the presence of a test compound, and detecting a change, for example, an increase or a diminution in a cellular

response of the second portion of the cells as compared to the first portion of the cells. The change in cellular response is shown as a measurable change extracellular acidification rate. Antagonists and agonists, for ZSMF-7 polypeptide, can be rapidly identified using this method.

Moreover, ZSMF-7 can be used to identify cells, tissues, or cell lines which respond to a ZSMF-7-stimulated pathway. The microphysiometer, described above, can be used to rapidly identify ligand-responsive cells, such as cells responsive to ZSMF-7 of the present invention. Cells can be cultured in the presence or absence of ZSMF-7 polypeptide. Those cells which elicit a measurable change in extracellular acidification in the presence of ZSMF-7 are responsive to ZSMF-7. Such cell lines, can be used to identify antagonists and agonists of ZSMF-7 polypeptide as described above.

ZSMF-7 polypeptides can also be used to identify inhibitors (antagonists) of its activity. ZSMF-7 antagonists include anti-ZSMF-7 antibodies and soluble ZSMF-7 receptors, as well as other peptidic and non-peptidic agents (including ribozymes). Test compounds are added to the assays disclosed herein to identify compounds that inhibit the activity of ZSMF-7. In addition to those assays disclosed herein, samples can be tested for inhibition of ZSMF-7 activity within a variety of assays designed to measure receptor binding or the stimulation/inhibition of ZSMF-7-dependent cellular responses. For example, ZSMF-7-responsive cell lines can be transfected with a reporter gene construct that is responsive to a ZDMF-7-stimulated cellular pathway. Reporter gene constructs of this type are known in the art, and will generally comprise a ZSMF-7-DNA response element operably linked to a gene encoding an assayable protein, such as luciferase. DNA response elements can include, but are not limited to, cyclic AMP response elements (CRE),

hormone response elements (HRE) insulin response element (IRE) (Nasrin et al., Proc. Natl. Acad. Sci. USA 87:5273-7, 1990) and serum response elements (SRE) (Shaw et al. Cell 56: 563-72, 1989). Cyclic AMP response elements are reviewed in Roestler et al., J. Biol. Chem. 263 (19):9063-6; 1988 and Habener, Molec. Endocrinol. 4 (8):1087-94; 1990. Hormone response elements are reviewed in Beato, Cell 56:335-44; 1989. Candidate compounds, solutions, mixtures or extracts are tested for the ability to inhibit the activity of ZSMF-7 on the target cells as evidenced by a decrease in ZSMF-7 stimulation of reporter gene expression. Assays of this type will detect compounds that directly block ZSMF-7 binding to cell-surface receptors, as well as compounds that block processes in the cellular pathway subsequent to receptor-ligand binding. In the alternative, compounds or other samples can be tested for direct blocking of ZSMF-7 binding to receptor using ZSMF-7 tagged with a detectable label (e.g., ^{125}I, biotin, horseradish peroxidase, FITC, and the like). Within assays of this type, the ability of a test sample to inhibit the binding of labeled ZSMF-7 to the receptor is indicative of inhibitory activity, which can be confirmed through secondary assays. Receptors used within binding assays may be cellular receptors or isolated, immobilized receptors.

ZSMF-7 antagonists would find use to modulate or down regulate one or more detrimental biological processes in cells, tissues and/or biological fluids, such as over-responsiveness, unregulated or inappropriate growth, and inflammation or allergic reaction. ZSMF-7 antagonists would have beneficial therapeutic effect in diseases where the inhibition of activation of certain B lymphocytes and/or T cells would be effective. In particular, such diseases would include autoimmune diseases, such as multiple sclerosis, insulin-dependent diabetes and systemic lupus erythematosus. Also, benefit would be derived from using ZSMF-7 antagonists for chronic inflammatory and

infective diseases. Antagonists could be used to dampen or inactivate ZSMF-7 during activated immune response.

The activity of semaphorin polypeptides, agonists, antagonists and antibodies of the present invention can be measured, and compounds screened to identify agonists and antagonists, using a variety of assays, such as assays that measure axon guidance and growth. Of particular interest are assays that indicate changes in neuron growth patterns, see for example, Hastings, WIPO Patent Application No:97/29189 and Walter et al., Development 101:685-96, 1987. Assays to measure the effects of semaphorins on neuron growth are well known in the art. For example, the C assay (see for example, Raper and Kapfhammer, Neuron 4:21-9, 1990 and Luo et al., Cell 75:217-27, 1993), can be used to determine collapsing activity semaphorins on growing neurons. Other methods which assess semaphorin induced inhibition of neurite extension or divert such extension are also known, see Goodman, Annu. Rev. Neurosci. 19:341-77, 1996. Conditioned media from cells expressing a semaphorin, semaphorin agonist or semaphorin antagonist, or aggregates of such cells, can by placed in a gel matrix near suitable neural cells, such as dorsal root ganglia (DRG) or sympathetic ganglia explants, which have been cocultured with nerve growth factor. Compared to control cells, semaphorin-induced changes in neuron growth can be measured (see for example, Messersmith et al., Neuron 14:949-59, 1995; Puschel et al., Neuron 14:941-8, 1995). Likewise neurite outgrowth can be measured using neuronal cell suspensions grown in the presence of molecules of the present invention see for example, O'Shea et al., Neuron 7:231-7, 1991 and DeFreitas et al., Neuron 15:333-43, 1995.

Also available are assay systems that use a ligand-binding receptor (or an antibody, one member of a complement/anti-complement pair) or a binding fragment thereof, and a commercially available biosensor instrument (BIAcore™, Pharmacia Biosensor, Piscataway, NJ). As used

herein, "complement/anti-complement pair" denotes non-identical moieties that form a non-covalently associated, stable pair under appropriate conditions. For instance, biotin and avidin (or streptavidin) are prototypical members of a complement/anti-complement pair. Other exemplary complement/anti-complement pairs include receptor/ligand pairs, antibody/antigen (or hapten or epitope) pairs, sense/antisense polynucleotide pairs, and the like. Where subsequent dissociation of the complement/anti-complement pair is desirable, the complement/anti-complement pair preferably has a binding affinity of $<10^9$ M^{-1}. Such receptor, antibody, member of a complement/anti-complement pair or fragment is immobilized onto the surface of a receptor chip. Use of this instrument is disclosed by Karlsson, J. Immunol. Methods 145:229-40, 1991 and Cunningham and Wells, J. Mol. Biol. 234:554-63, 1993. A receptor, antibody, member or fragment is covalently attached, using amine or sulfhydryl chemistry, to dextran fibers that are attached to gold film within the flow cell. A test sample is passed through the cell. If a ligand, epitope, or opposite member of the complement/anti-complement pair is present in the sample, it will bind to the immobilized receptor, antibody or member, respectively, causing a change in the refractive index of the medium, which is detected as a change in surface plasmon resonance of the gold film. This system allows the determination of on- and off-rates, from which binding affinity can be calculated, and assessment of stoichiometry of binding. Ligand-binding receptor polypeptides can also be used within other assay systems known in the art. Such systems include Scatchard analysis for determination of binding affinity (see, Scatchard, Ann. NY Acad. Sci. 51: 660-72, 1949) and calorimetric assays (Cunningham et al., Science 253:545-8, 1991; Cunningham et al., Science 245:821-5, 1991).

Proteins of the present invention may also be assayed using viral delivery systems. Exemplary viruses

for this purpose include adenovirus, herpesvirus, vaccinia virus and adeno-associated virus (AAV). Adenovirus, a double-stranded DNA virus, is currently the best studied gene transfer vector for delivery of heterologous nucleic acid (for a review, see T.C. Becker et al., Meth. Cell Biol. 43:161-89, 1994; and J.T. Douglas and D.T. Curiel, Science & Medicine 4:44-53, 1997). The adenovirus system offers several advantages: adenovirus can (i) accommodate relatively large DNA inserts; (ii) be grown to high-titer; (iii) infect a broad range of mammalian cell types; and (iv) be used with a large number of available vectors containing different promoters. Also, because adenoviruses are stable in the bloodstream, they can be administered by intravenous injection.

By deleting portions of the adenovirus genome, larger inserts (up to 7 kb) of heterologous DNA can be accommodated. These inserts can be incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. In an exemplary system, the essential E1 gene has been deleted from the viral vector, and the virus will not replicate unless the E1 gene is provided by the host cell (the human 293 cell line is exemplary). When intravenously administered to intact animals, adenovirus primarily targets the liver. If the adenoviral delivery system has an E1 gene deletion, the virus cannot replicate in the host cells. However, the host's tissue (e.g., liver) will express and process (and, if a secretory signal sequence is present, secrete) the heterologous protein. Secreted proteins will enter the circulation in the highly vascularized liver, and effects on the infected animal can be determined.

The adenovirus system can also be used for protein production *in vitro*. By culturing adenovirus-infected non-293 cells under conditions where the cells are not rapidly dividing, the cells can produce proteins for extended periods of time. For instance, BHK cells are grown to confluence in cell factories, then exposed to the

adenoviral vector encoding the secreted protein of interest. The cells are then grown under serum-free conditions, which allows infected cells to survive for several weeks without significant cell division. Alternatively, adenovirus vector infected 293S cells can be grown in suspension culture at relatively high cell density to produce significant amounts of protein (see Garnier et al., Cytotechnol. 15:145-55, 1994). With either protocol, an expressed, secreted heterologous protein can be repeatedly isolated from the cell culture supernatant. Within the infected 293S cell production protocol, non-secreted proteins may also be effectively obtained.

ZSMF-7 polypeptides can also be used to prepare antibodies that specifically bind to ZSMF-7 polypeptides. As used herein, the term "antibodies" includes polyclonal antibodies, monoclonal antibodies, antigen-binding fragments thereof such as $F(ab')_2$ and Fab fragments, single chain antibodies, and the like, including genetically engineered antibodies.

For particular uses, it may be desirable to prepare fragments of anti-ZSMF-7 antibodies. Such antibody fragments can be obtained, for example, by proteolytic hydrolysis of the antibody. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. As an illustration, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted $F(ab')_2$. This fragment can be further cleaved using a thiol reducing agent to produce 3.5S Fab' monovalent fragments. Optionally, the cleavage reaction can be performed using a blocking group for the sulfhydryl groups that result from cleavage of disulfide linkages. As an alternative, an enzymatic cleavage using pepsin produces two monovalent Fab fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patent No. 4,331,647, Nisonoff et al., Arch Biochem. Biophys. 89:230, 1960, Porter, Biochem. J. 73:119, 1959,

Edelman et al., in Methods in Enzymology Vol. 1, page 422 (Academic Press 1967), and by Coligan, ibid.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

For example, Fv fragments comprise an association of V_H and V_L chains. This association can be noncovalent, as described by Inbar et al., Proc. Natl. Acad. Sci. USA 69:2659, 1972. Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as gluteraldehyde (see, for example, Sandhu, Crit. Rev. Biotech. 12:437, 1992).

The Fv fragments may comprise V_H and V_L chains which are connected by a peptide linker. These single-chain antigen binding proteins (scFv) are prepared by constructing a structural gene comprising DNA sequences encoding the V_H and V_L domains which are connected by an oligonucleotide. The structural gene is inserted into an expression vector which is subsequently introduced into a host cell, such as *E. coli*. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing scFvs are described, for example, by Whitlow et al., Methods: A Companion to Methods in Enzymology 2:97, 1991, also see, Bird et al., Science 242:423, 1988, Ladner et al., U.S. Patent No. 4,946,778, Pack et al., Bio/Technology 11:1271, 1993, and Sandhu, *supra*.

As an illustration, a scFV can be obtained by exposing lymphocytes to ZSMF-7 polypeptide *in vitro*, and selecting antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled ZSMF-7 protein or peptide). Genes encoding polypeptides having potential ZSMF-7 polypeptide binding domains can be obtained by screening random peptide

libraries displayed on phage (phage display) or on bacteria, such as *E. coli.* Nucleotide sequences encoding the polypeptides can be obtained in a number of ways, such as through random mutagenesis and random polynucleotide synthesis. These random peptide display libraries can be used to screen for peptides which interact with a known target which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner et al., U.S. Patent No. 5,223,409, Ladner et al., U.S. Patent No. 4,946,778, Ladner et al., U.S. Patent No. 5,403,484, Ladner et al., U.S. Patent No. 5,571,698, and Kay et al., <u>Phage Display of Peptides and Proteins</u> (Academic Press, Inc. 1996)) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from Clontech (Palo Alto, CA), Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA), and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Random peptide display libraries can be screened using the ZSMF-7 sequences disclosed herein to identify proteins which bind to ZSMF-7.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells (see, for example, Larrick et al., <u>Methods: A Companion to Methods in Enzymology</u> <u>2</u>:106, 1991), Courtenay-Luck, "Genetic Manipulation of Monoclonal Antibodies," in <u>Monoclonal Antibodies: Production, Engineering and Clinical Application</u>, Ritter et al. (eds.), page 166 (Cambridge University Press 1995), and Ward et al., "Genetic Manipulation and Expression of Antibodies,"

in Monoclonal Antibodies: Principles and Applications, Birch et al., (eds.), page 137 (Wiley-Liss, Inc. 1995)).

Non-human antibodies can be humanized by grafting non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans is reduced. One skilled in the art can generate humanized antibodies with specific and different constant domains (i.e., different Ig subclasses) to facilitate or inhibit various immune functions associated with particular antibody constant domains. Moreover, human antibodies can be produced in transgenic, non-human animals that have been engineered to contain human immunoglobulin genes as disclosed in WIPO Publication WO 98/24893. It is preferred that the endogenous immunoglobulin genes in these animals be inactivated or eliminated, such as by homologous recombination.

Alternative techniques for generating or selecting antibodies useful herein include *in vitro* exposure of lymphocytes to ZSMF-7 polypeptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled ZSMF-7 polypeptide).

Antibodies are determined to be specifically binding if: 1) they exhibit a threshold level of binding activity, and/or 2) they do not significantly cross-react with related polypeptide molecules. Antibodies herein specifically bind if they bind to a human ZSMF-7 polypeptide, peptide or epitope with a binding affinity (K_a) of 10^6 mol^{-1} or greater, preferably 10^7 mol^{-1} or

greater, more preferably 10^8 mol^{-1} or greater, and most preferably 10^9 mol^{-1} or greater. The binding affinity of an antibody can be readily determined by one of ordinary skill in the art, for example, by Scatchard analysis (Scatchard, ibid.). Antibodies of the current invention do not significantly cross-react with related polypeptide molecules, for example, if they detect ZSMF-7 but not known related polypeptides using a standard Western blot analysis (Ausubel et al., ibid.). Examples of known related polypeptides are orthologs; proteins from the same species that are members of a protein family such as other known semaphorins (Sema A-Sema G, Sema IV and CD 100); mutant semaphorin polypeptides; and non-human semaphorins (G Sema I, D Sema I and II and T Sema I). Moreover, antibodies may be "screened against" known related polypeptides to isolate a population that specifically binds to the inventive polypeptides. For example, antibodies raised to ZSMF-7 are adsorbed to related polypeptides adhered to insoluble matrix; antibodies specific to ZSMF-7 will flow through the matrix under the proper buffer conditions. Such screening allows isolation of polyclonal and monoclonal antibodies non-crossreactive to closely related polypeptides (Antibodies: A Laboratory Manual, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988; Current Protocols in Immunology, Cooligan, et al. (eds.), National Institutes of Health, John Wiley and Sons, Inc., 1995). Screening and isolation of specific antibodies is well known in the art (see, Fundamental Immunology, Paul (eds.), Raven Press, 1993; Getzoff et al., Adv. in Immunol. 43:1-98, 1988; Monoclonal Antibodies: Principles and Practice, Goding, J.W. (eds.), Academic Press Ltd., 1996; Benjamin et al., Ann. Rev. Immunol. 2:67-101, 1984).

Methods for preparing polyclonal and monoclonal antibodies are well known in the art (see for example, Hurrell, Ed., Monoclonal Hybridoma Antibodies: Techniques and Applications, CRC Press, Inc., Boca Raton, FL, 1982). As would be evident to one of ordinary skill in the art,

polyclonal antibodies can be generated from a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, and rats. The immunogenicity of a ZSMF-7 polypeptide can be increased through the use of an adjuvant such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of a ZSMF-7 polypeptide or a portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof. If the polypeptide portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

Alternative techniques for generating or selecting antibodies useful herein include *in vitro* exposure of lymphocytes to ZSMF-7 polypeptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled ZSMF-7 polypeptide).

Polyclonal anti-idiotype antibodies can be prepared by immunizing animals with anti-ZSMF-7 antibodies or antibody fragments, using standard techniques. See, for example, Green et al., "Production of Polyclonal Antisera," in Methods In Molecular Biology: Immunochemical Protocols, Manson (ed.), pages 1-12 (Humana Press 1992). Also, see Coligan, ibid. at pages 2.4.1-2.4.7. Alternatively, monoclonal anti-idiotype antibodies can be prepared using anti-ZSMF-7 antibodies or antibody fragments as immunogens with the techniques, described above. As another alternative, humanized anti-idiotype antibodies or subhuman primate anti-idiotype antibodies can be prepared using the above-described techniques. Methods for producing anti-idiotype antibodies are described, for example, by Irie, U.S. Patent No. 5,208,146, Greene, et. al., U.S. Patent No.

5,637,677, and Varthakavi and Minocha, J. Gen. Virol. 77:1875, 1996.

A variety of assays known to those skilled in the art can be utilized to detect antibodies that specifically bind to ZSMF-7 polypeptides. Exemplary assays are described in detail in Antibodies: A Laboratory Manual, Harlow and Lane (Eds.), Cold Spring Harbor Laboratory Press, 1988. Representative examples of such assays include concurrent immunoelectrophoresis, radio-immunoassays, radio-immunoprecipitations, enzyme-linked immunosorbent assays (ELISA), dot blot assays, Western blot assays, inhibition or competition assays, and sandwich assays. In addition, antibodies can be screened for binding to wild-type versus mutant ZSMF-7 protein or peptides.

Antibodies to ZSMF-7 can be used for affinity purification of ZSMF-7 polypeptides; within diagnostic assays for determining circulating levels of ZSMF-7 polypeptides; for detecting or quantitating soluble ZSMF-7 polypeptide as a marker of underlying pathology or disease; for immunolocalization within whole animals or tissue sections, including immunodiagnostic applications; for immunohistochemistry; and as antagonists to block protein activity *in vitro* and *in vivo*. Antibodies to ZSMF-7 can also be used for tagging cells that express ZSMF-7; for affinity purification of ZSMF-7 polypeptides; in analytical methods employing FACS; for screening expression libraries; and for generating anti-idiotypic antibodies. Antibodies can be linked to other compounds, including therapeutic and diagnostic agents, using known methods to provide for targeting of those compounds to cells expressing receptors for ZSMF-7. For certain applications, including *in vitro* and *in vivo* diagnostic uses, it is advantageous to employ labeled antibodies. Suitable direct tags or labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, chemiluminescent markers, magnetic particles and the like; indirect tags or labels may feature

use of biotin-avidin or other complement/anti-complement pairs as intermediates. Antibodies of the present invention can also be directly or indirectly conjugated to drugs, toxins, radionuclides and the like, and these conjugates used for *in vivo* diagnostic or therapeutic applications.

Genes encoding polypeptides having potential ZSMF-7 polypeptide binding domains can be obtained by screening random peptide libraries displayed on phage (phage display) or on bacteria, such as *E. coli*. Nucleotide sequences encoding the polypeptides can be obtained in a number of ways, such as through random mutagenesis and random polynucleotide synthesis. These random peptide display libraries can be used to screen for peptides which interact with a known target which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner et al., US Patent NO. 5,223,409; Ladner et al., US Patent NO. 4,946,778; Ladner et al., US Patent NO. 5,403,484 and Ladner et al., US Patent NO. 5,571,698) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from Clontech (Palo Alto, CA), Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA) and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Random peptide display libraries can be screened using the ZSMF-7 sequences disclosed herein to identify proteins which bind to ZSMF-7. These "binding proteins" which interact with ZSMF-7 polypeptides can be used for tagging cells; for isolating homolog polypeptides by affinity purification; they can be directly or indirectly conjugated to drugs, toxins, radionuclides and the like. These binding proteins can also be used in analytical methods such as for screening expression libraries and neutralizing activity. The binding proteins can also be used for diagnostic assays

for determining circulating levels of polypeptides; for detecting or quantitating soluble polypeptides as marker of underlying pathology or disease. These binding proteins can also act as ZSMF-7 "antagonists" to block ZSMF-7 binding and signal transduction *in vitro* and *in vivo*. These anti-ZSMF-7 binding proteins would be useful for inhibiting ZSMF-7 binding.

ZSMF-7 polypeptides and polynucleotides may be used within diagnostic systems. Antibodies or other agents that specifically bind to ZSMF-7 may be used to detect the presence of circulating ligand or receptor polypeptides. Such detection methods are well known in the art and include, for example, enzyme-linked immunosorbent assay (ELISA) and radioimmunoassay. Immunohistochemically labeled ZSMF-7 antibodies can be used to detect ZSMF-7 receptor and/or ligands in tissue samples and identify ZSMF-7 receptors. ZSMF-7 levels can also be monitored by such methods as RT-PCR, where ZSMF-7 mRNA can be detected and quantified. The information derived from such detection methods would provide insight into the significance of ZSMF-7 polypeptides in various diseases and biological processes, and as a would serve as diagnostic tools for diseases for which altered levels of ZSMF-7 are significant.

Nucleic acid molecules disclosed herein can be used to detect the expression of a ZSMF-7 gene in a biological sample. Such probe molecules include double-stranded nucleic acid molecules comprising the nucleotide sequences of SEQ ID NOs:1 or 5, or fragments thereof, as well as single-stranded nucleic acid molecules having the complement of the nucleotide sequences of SEQ ID NOs:1 or 5, or a fragment thereof. Probe molecules may be DNA, RNA, oligonucleotides, and the like.

As an illustration, suitable probes include nucleic acid molecules that bind with a portion of a ZSMF-7 domain or motif, such as the ZSMF-7 semaphorin domain

(nucleotides 69-541 of SEQ ID NO:1 or nucleotides 205-1623 of SEQ ID NO:5). Other probes include those to the Ig-like domain.

In a basic assay, a single-stranded probe molecule is incubated with RNA, isolated from a biological sample, under conditions of temperature and ionic strength that promote base pairing between the probe and target ZSMF-7 RNA species. After separating unbound probe from hybridized molecules, the amount of hybrids is detected.

Well-established hybridization methods of RNA detection include northern analysis and dot/slot blot hybridization, see, for example, Ausubel ibid. and Wu et al. (eds.), "Analysis of Gene Expression at the RNA Level," in Methods in Gene Biotechnology, pages 225-239 (CRC Press, Inc. 1997), and methods described herein. Nucleic acid probes can be detectably labeled with radioisotopes such as ^{32}P or ^{35}S. Alternatively, ZSMF-7 RNA can be detected with a nonradioactive hybridization method (see, for example, Isaac (ed.), Protocols for Nucleic Acid Analysis by Nonradioactive Probes, Humana Press, Inc., 1993). Typically, nonradioactive detection is achieved by enzymatic conversion of chromogenic or chemiluminescent substrates. Illustrative non-radioactive moieties include biotin, fluorescein, and digoxigenin.

ZSMF-7 oligonucleotide probes are also useful for *in vivo* diagnosis. As an illustration, ^{18}F-labeled oligonucleotides can be administered to a subject and visualized by positron emission tomography (Tavitian et al., Nature Medicine 4:467, 1998).

Numerous diagnostic procedures take advantage of the polymerase chain reaction (PCR) to increase sensitivity of detection methods. Standard techniques for performing PCR are well-known (see, generally, Mathew (ed.), Protocols in Human Molecular Genetics (Humana Press, Inc. 1991), White (ed.), PCR Protocols: Current Methods and Applications (Humana Press, Inc. 1993), Cotter (ed.), Molecular Diagnosis of Cancer (Humana Press, Inc. 1996),

Hanausek and Walaszek (eds.), Tumor Marker Protocols (Humana Press, Inc. 1998), Lo (ed.), Clinical Applications of PCR (Humana Press, Inc. 1998), and Meltzer (ed.), PCR in Bioanalysis (Humana Press, Inc. 1998)). PCR primers can be designed to amplify a sequence encoding a particular ZSMF-7 domain or motif, such as the ZSMF-7 semaphorin domain (nucleotides 69-541 of SEQ ID NO:1 or nucleotides 205-1623 of SEQ ID NO:5).

One variation of PCR for diagnostic assays is reverse transcriptase-PCR (RT-PCR). In the RT-PCR technique, RNA is isolated from a biological sample, reverse transcribed to cDNA, and the cDNA is incubated with ZSMF-7 primers (see, for example, Wu et al. (eds.), "Rapid Isolation of Specific cDNAs or Genes by PCR," in Methods in Gene Biotechnology, CRC Press, Inc., pages 15-28, 1997). PCR is then performed and the products are analyzed using standard techniques.

As an illustration, RNA is isolated from biological sample using, for example, the guanidinium-thiocyanate cell lysis procedure described herein. Alternatively, a solid-phase technique can be used to isolate mRNA from a cell lysate. A reverse transcription reaction can be primed with the isolated RNA using random oligonucleotides, short homopolymers of dT, or ZSMF-7 anti-sense oligomers. Oligo-dT primers offer the advantage that various mRNA nucleotide sequences are amplified that can provide control target sequences. ZSMF-7 sequences are amplified by the polymerase chain reaction using two flanking oligonucleotide primers that are typically at least 5 bases in length.

PCR amplification products can be detected using a variety of approaches. For example, PCR products can be fractionated by gel electrophoresis, and visualized by ethidium bromide staining. Alternatively, fractionated PCR products can be transferred to a membrane, hybridized with a detectably-labeled ZSMF-7 probe, and examined by

autoradiography. Additional alternative approaches include the use of digoxigenin-labeled deoxyribonucleic acid triphosphates to provide chemiluminescence detection, and the C-TRAK colorimetric assay.

Another approach is real time quantitative PCR (Perkin-Elmer Cetus, Norwalk, Ct.). A fluorogenic probe, consisting of an oligonucleotide with both a reporter and a quencher dye attached, anneals specifically between the forward and reverse primers. Using the 5′ endonuclease activity of Taq DNA polymerase, the reporter dye is separated from the quencher dye and a sequence-specific signal is generated and increases as amplification increases. The fluorescence intensity can be continuously monitored and quantified during the PCR reaction.

Another approach for detection of ZSMF-7 expression is cycling probe technology (CPT), in which a single-stranded DNA target binds with an excess of DNA-RNA-DNA chimeric probe to form a complex, the RNA portion is cleaved with RNase H, and the presence of cleaved chimeric probe is detected (see, for example, Beggs et al., J. Clin. Microbiol. 34:2985, 1996 and Bekkaoui et al., Biotechniques 20:240, 1996). Alternative methods for detection of ZSMF-7 sequences can utilize approaches such as nucleic acid sequence-based amplification (NASBA), cooperative amplification of templates by cross-hybridization (CATCH), and the ligase chain reaction (LCR) (see, for example, Marshall et al., U.S. Patent No. 5,686,272 (1997), Dyer et al., J. Virol. Methods 60:161, 1996; Ehricht et al., Eur. J. Biochem. 243:358, 1997 and Chadwick et al., J. Virol. Methods 70:59, 1998). Other standard methods are known to those of skill in the art.

ZSMF-7 probes and primers can also be used to detect and to localize ZSMF-7 gene expression in tissue samples. Methods for such *in situ* hybridization are well-known to those of skill in the art (see, for example, Choo (ed.), In Situ Hybridization Protocols, Humana Press, Inc., 1994; Wu et al. (eds.), "Analysis of Cellular DNA or

Abundance of mRNA by Radioactive *In Situ* Hybridization (RISH)," in Methods in Gene Biotechnology, CRC Press, Inc., pages 259-278, 1997 and Wu et al. (eds.), "Localization of DNA or Abundance of mRNA by Fluorescence *In Situ* Hybridization (RISH)," in Methods in Gene Biotechnology, CRC Press, Inc., pages 279-289, 1997).

Various additional diagnostic approaches are well-known to those of skill in the art (see, for example, Mathew (ed.), Protocols in Human Molecular Genetics Humana Press, Inc., 1991; Coleman and Tsongalis, Molecular Diagnostics, Humana Press, Inc., 1996 and Elles, Molecular Diagnosis of Genetic Diseases, Humana Press, Inc., 1996).

The ZSMF-7 polynucleotides and/or polypeptides disclosed herein can be useful as therapeutics, wherein ZSMF-7 agonists and antagonists could modulate one or more biological processes in cells, tissues and/or biological fluids. ZSMF-7 antagonists provided by the invention, bind to ZSMF-7 polypeptides or, alternatively, to a receptor to which ZSMF-7 polypeptides bind, thereby inhibiting or eliminating the function of ZSMF-7. Such ZSMF-7 antagonists would include antibodies; oligonucleotides which bind either to the ZSMF-7 polypeptide or to its ligand; natural or synthetic analogs of ZSMF-7 ligands which retain the ability to bind the receptor but do not result in either ligand or receptor signaling. Such analogs could be peptides or peptide-like compounds. Natural or synthetic small molecules which bind to ZSMF-7 polypeptides and prevent signaling are also contemplated as antagonists. As such, ZSMF-7 antagonists would be useful as therapeutics for treating certain disorders where blocking signal from either a ZSMF-7 receptor or ligand would be beneficial.

The invention also provides nucleic acid-based therapeutic treatment. If a mammal lacks or has a mutated ZSMF-7 gene, the ZSMF-7 gene can be introduced into the cells of the mammal. Using such methods, cells altered to

express the nerve growth factor neurotrophin-3 (NT-3) were grafted to a rat model for spinal injury and stimulated axon regrowth at the lesion site and the rats thus treated recovered some ability to walk (Grill et al., J. Neuroscience 17:5560-72, 1997). In one embodiment, a gene encoding a ZSMF-7 polypeptide is introduced *in vivo* in a viral vector. Such vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. A defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt et al., Molec. Cell. Neurosci. 2:320-30, 1991), an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet et al. (J. Clin. Invest. 90:626-30, 1992), and a defective adeno-associated virus vector (Samulski et al., J. Virol. 61:3096-101, 1987; Samulski et al., J. Virol. 63:3822-8, 1989).

In another embodiment, the gene can be introduced in a retroviral vector, *e.g.*, as described in Anderson et al., U.S. Patent No. 5,399,346; Mann et al., Cell 33:153, 1983; Temin et al., U.S. Patent No. 4,650,764; Temin et al., U.S. Patent No. 4,980,289; Markowitz et al., J. Virol. 62:1120, 1988; Temin et al., U.S. Patent No. 5,124,263; Dougherty et al., WIPO Publication WO 95/07358; and Kuo et al., Blood 82:845-52, 1993.

Alternatively, the vector can be introduced by lipofection *in vivo* using liposomes. Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7, 1987; and Mackey et

al., Proc. Natl. Acad. Sci. USA 85:8027-31, 1988). The use of lipofection to introduce exogenous genes into specific organs *in vivo* has certain practical advantages. Molecular targeting of liposomes to specific cells represents one area of benefit. It is clear that directing transfection to particular cells represents one area of benefit. It is clear that directing transfection to particular cell types would be particularly advantageous in a tissue with cellular heterogeneity, such as the pancreas, liver, kidney, and brain. Lipids may be chemically coupled to other molecules for the purpose of targeting. Targeted peptides, e.g., hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

It is possible to remove the cells from the body and introduce the vector as a naked DNA plasmid and then re-implant the transformed cells into the body. Naked DNA vector for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun or use of a DNA vector transporter (see, for example, Wu et al., J. Biol. Chem. 267:963-7, 1992; Wu et al., J. Biol. Chem. 263:14621-4, 1988).

Another aspect of the present invention involves antisense polynucleotide compositions that are complementary to a segment of the polynucleotide set forth in SEQ ID NO:1. Such synthetic antisense oligonucleotides are designed to bind to mRNA encoding ZSMF-7 polypeptides and to inhibit translation of such mRNA. Such antisense oligonucleotides are used to inhibit expression of ZSMF-7 polypeptide-encoding genes in cell culture or in a subject.

Transgenic mice, engineered to express the ZSMF-7 gene, and mice that exhibit a complete absence of ZSMF-7 gene function, referred to as "knockout mice" (Snouwaert et al., Science 257:1083, 1992), may also be generated (Lowell et al., Nature 366:740-2, 1993). These mice may be

employed to study the ZSMF-7 gene and the protein encoded thereby in an *in vivo* system.

For pharmaceutical use, the proteins of the present invention are formulated for parenteral, particularly intravenous or subcutaneous, delivery according to conventional methods. Intravenous administration will be by bolus injection or infusion over a typical period of one to several hours. In general, pharmaceutical formulations will include a ZSMF-7 polypeptide in combination with a pharmaceutically acceptable vehicle, such as saline, buffered saline, 5% dextrose in water or the like. Formulations may further include one or more excipients, preservatives, solubilizers, buffering agents, albumin to prevent protein loss on vial surfaces, etc. Methods of formulation are well known in the art and are disclosed, for example, in Remington: The Science and Practice of Pharmacy, Gennaro, ed., Mack Publishing Co., Easton, PA, 19th ed., 1995. Determination of dose is within the level of ordinary skill in the art.

The invention is further illustrated by the following non-limiting examples.

EXAMPLES

Example 1
Identification of ZSMF-7

Novel ZSMF-7 encoding polynucleotides and polypeptides of the present invention were initially identified by querying an EST database for sequences homologous to conserved motifs within the semaphorin family. Expressed sequence tags (ESTs) from human retina, human placenta and human fibroblasts cDNA libraries that corresponded the 5' end of the gene were identified.

To obtain the complete cDNA sequence of ZSMF-7, a human testis library was screened. The construction of the cDNA libraries is known in the art and such libraries may be purchased from commercial suppliers such as Clontech Laboratories, Inc. (Palo Alto, CA). The library was plated in pools of 5000 colonies/pool. Plasmid DNA was prepared from the plated bacteria using a Qiagen® plasmid purification column (Qiagen, Inc., Chatsworth, CA) according to the manufacturer's instructions. DNA from these pools were combined into larger pools. Oligonucleotides ZC16,189 (SEQ ID NO:24) and ZC16188 (SEQ ID NO:25) were designed from an incomplete clone obtained from a human placenta library for use as PCR primers. Using the pooled human testis library DNA as a template, amplification was carried out as follows: 35 cycles of 94°C for 30 seconds, 58°C for 30 seconds and 72°C for 30 seconds. Positive clones were identified by the presence of a 583 bp PCR fragment (SEQ ID NO:26). Two pools of 5000 colonies were found to contain this fragment. These pools were used to transform *E. coli* which were plated to agar. The colonies were transferred to nylon membrane and probed with the 583 bp PCR fragment (SEQ ID NO:26). The fragment was gel purified using a Qiaquick kit (Qiagen, Inc., Chatsworth, CA) and radioactively labeled using the random priming MULTIPRIME DNA labeling system (Amersham, Arlington

Heights, IL), according to the manufacturer's specifications. The probe was purified using a NUCTRAP push column (Stratagene, La Jolla, CA). ExpressHyb (Clontech) solution was used for prehybridization and as a hybridizing solution for the colony lifts. The filters were hybridized with the labeled probe at 65°C, overnight, and then washed with an SSC/SDS buffer under appropriately stringent conditions and positive colonies detected upon exposure to film. Plasmid DNA from colonies producing signal was then isolated and submitted for sequence analysis. The plasmid DNA from a positive colony was used as template and oligos ZC694 (SEQ ID NO:8) and ZC2681 (SEQ ID NO:22) to the vector were used as sequencing primers. Oligonucleotides ZC16820 (SEQ ID NO:9), ZC16087 (SEQ ID NO:10), ZC16818 (SEQ ID NO:11), ZC15394 (SEQ ID NO:12), ZC16819 (SEQ ID NO:13), ZC16460 (SEQ ID NO:14), ZC16548 (SEQ ID NO:15), ZC16807 (SEQ ID NO:16), ZC16806 (SEQ ID NO:17), ZC16667 (SEQ ID NO:18), ZC16729 (SEQ ID NO:19), ZC16728 (SEQ ID NO:20) and ZC16666 (SEQ ID NO:21) were used to complete the sequence. Sequencing reactions were carried out in a Hybaid OmniGene Temperature Cycling System (National Labnet Co., Woodbridge, NY). Sequencher™ 3.0 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI) was used for data analysis. The resulting 3,377 bp sequence is disclosed in SEQ ID NO:1.

Example 2
Tissue Distribution

Human Multiple Tissue Northern Blots (MTN I, MTN II, and MTN III; Clontech) were probed to determine the tissue distribution of human ZSMF-7 expression. An approximately 234 bp probe (SEQ ID NO:4) was amplified from a human retina derived Marathon™-ready cDNA library. Oligonucleotide primers ZC14298 (SEQ ID NO:27) and ZC14299 (SEQ ID NO:28) were designed based on an EST sequence. The Marathon™-ready cDNA library was prepared according to

manufacturer's instructions (Marathon™ cDNA Amplification Kit; Clontech) using human retina poly A+ RNA (Clontech). The probe was amplified in a polymerase chain reaction as follows: 1 cycle at 94°C for 1 minute; 35 cycles of 94°C for 30 seconds and 68°C for 1 minute 30 seconds, followed by 1 cycle at 72°C for 10 minutes. The resulting DNA fragment was electrophoresed on a 2% low melt agarose gel (SEA PLAQUE GTG low melt agarose, FMC Corp., Rockland, ME), the fragment was purified using the QIAquick™ method (Qiagen, Chatsworth, CA), and the sequence was confirmed by sequence analysis. The probe was radioactively labeled and purified as described herein. ExpressHyb™ (Clontech) solution was used for prehybridization and as a hybridizing solution for the Northern blots. Hybridization took place overnight at 65°C using 1.0 x 10^6 cpm/ml of labeled probe. The blots were then washed 4 times at room temperature in 2X SSC, 0.05% SDS followed by 2 washes at 50°C in 0.1X SSC, 0.01% SDS for 20 minutes each. A transcript of approximately 4.0 kb was seen in testis, spleen, spinal cord and placenta, a weak signal was detected in brain, thymus, ovary, lymph node and bone marrow.

Additional analysis was carried out on Northern blots made with poly(A) RNA from the human vascular cell lines HUVEC (human umbilical vein endothelial cells; Cascade Biologics, Inc., Portland, OR), HPAEC (human pulmonary artery endothelial cells; Cascade Biologics, Inc.), HAEC (human aortic endothelial cells; Cascade Biologics, Inc.), AoSMC (aortic smooth muscle cells; Clonetics, San Diego, CA), UASMC (umbilical artery smooth muscle cells; Clonetics), HISM (human intestinal smooth muscle cells; ATCC CRL 7130), SK-5 (human dermal fibroblast cells; obtained from Dr. Russell Ross, University of Washington, Seattle, WA), NHLF (normal human lung fibroblast cells; Clonetics), and NHDF-NEO (normal human dermal fibroblast-neonatal cells; Clonetics). The probe was prepared and labeled and prehybridization and hybridization were carried out essentially as disclosed above. The blots

were then washed at 50°C in 0.1X SSC, 0.05% SDS. A transcript of approximately 4.0 kb was seen in was seen in VASMC, AoSMC, SK-5, NHLF and NHDF-Neo cells. Signal intensity was highest in NHLF cells.

Additional analysis was carried out on Northern blots made with poly(A) RNA from K-562 cells (erythroid, ATCC CCL 243), HUT78 cells (T cell, ATCC TIB-161), Jurkat cells (T cell), DAUDI (Burkitt's human lymphoma, Clontech, Palo Alto, CA), RAJI (Burkitt's human lymphoma, Clontech) and HL60 (Monocyte). The probe preparation and hybridization were carried out as above. Two transcripts, approximately, ~4.5 and 4.0, were seen in DAUDI, RAJI, JRUKAT, HUT78 and HL60 cells. Signal intensity was highest in RAJI and JURKAT.

Additional analysis was carried out on Northern blots made with poly (A) RNA from $CD4^+$, $CD8^+$, $CD19^+$ and mixed lymphocyte reaction cells (CellPro, Bothell, WA) using probes and hybridization conditions described above. A transcript of approximately 4.0 kb was seen in the mixed lymphocytes and CD19+ cells. Signal intensity was highest in the mixed lymphocyte cells.

Additional analysis was carried out on Human Brain Multiple Tissue Northern Blots II and III (Clontech) using the probe and hybridization conditions described above. A transcript of 4.0 kb was seen in all tissue tested.

Example 3
Chromosomal Assignment and Placement of ZSMF-7

ZSMF-7 was mapped to chromosome 15 using the commercially available GeneBridge 4 Radiation Hybrid Panel (Research Genetics, Inc., Huntsville, AL). The GeneBridge 4 Radiation Hybrid Panel contains PCRable DNAs from each of 93 radiation hybrid clones, plus two control DNAs (the HFL donor and the A23 recipient). A publicly available WWW server (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl) allows mapping relative to the Whitehead

Institute/MIT Center for Genome Research's radiation hybrid map of the human genome (the "WICGR" radiation hybrid map) which was constructed with the GeneBridge 4 Radiation Hybrid Panel.

For the mapping of ZSMF-7 with the GeneBridge 4 RH Panel, 20 μl reactions were set up in a 96-well microtiter plate (Stratagene, La Jolla, CA) and used in a RoboCycler Gradient 96 thermal cycler (Stratagene). Each of the 95 PCR reactions consisted of 2 μl 10X KlenTaq PCR reaction buffer (Clontech), 1.6 μl dNTPs mix (2.5 mM each, PERKIN-ELMER, Foster City, CA), 1 μl sense primer, ZC 16086 (SEQ ID NO:6), 1 μl antisense primer, ZC 16,085 (SEQ ID NO:7), 2 μl *Redi*Load (Research Genetics, Inc.), 0.4 μl 50X Advantage KlenTaq Polymerase Mix (Clontech), 25 ng of DNA from an individual hybrid clone or control and ddH_2O for a total volume of 20 μl. The reactions were overlaid with an equal amount of mineral oil and sealed. The PCR cycler conditions were as follows: an initial 1 cycle 5 minute denaturation at 95°C, 35 cycles of a 1 minute denaturation at 95°C, 1 minute annealing at 66°C and 1.5 minute extension at 72°C, followed by a final 1 cycle extension of 7 minutes at 72°C. The reactions were separated by electrophoresis on a 2% agarose gel (Life Technologies, Gaithersburg, MD).

The results showed that ZSMF-7 maps 3.98 cR_3000 from the framework marker CHLC.GATA85D02 on the WICGR radiation hybrid map. Proximal and distal framework markers were CHLC.GATA85D02 and CHLC.GCT7C09, respectively. The use of surrounding markers positions ZSMF-7 in the 15q24.3 region on the integrated LDB chromosome 15 map (The Genetic Location Database, University of Southhampton, WWW server: http://cedar. genetics.soton.ac.uk/public_html/).

Example 4
ZSMF-7 Anti-peptide Antibodies

Polyclonal anti-peptide antibodies were prepared by immunizing two female New Zealand white rabbits and 5

mice with the peptide, huzsmf7-2 NIGSTKGSCLDKRDC ENYITLLERRSEGLLACGTNA (SEQ ID NO:35) from the N-terminal region of the semaphorin domain or huzsmf7-3 SINPAEPHKECPNPKPDKC (SEQ ID NO:36) from the C-terminal portion of the semaphorin domain. The peptides were synthesized using an Applied Biosystems Model 431A peptide synthesizer (Applied Biosystems, Inc., Foster City, CA) according to manufacturer's instructions. The peptides were then conjugated to the carrier protein maleimide-activated keyhole limpet hemocyanin (KLH). The rabbits were each given an initial intraperitoneal (ip) injection of 200 µg of peptide in Complete Freund's Adjuvant followed by booster ip injections of 100 µg peptide in Incomplete Freund's Adjuvant every three weeks. Seven to ten days after the administration of the second booster injection, the animals were bled and the serum was collected. The animals were then boosted and bled every three weeks.

The mice were each given an initial ip injection of 20 µg of peptide in Complete Freund's Adjuvant followed by booster ip injections of 10 µg peptide in Incomplete Freund's Adjuvant every two weeks. Seven to ten days after the administration of the second booster injection, the animals were bled and the serum was collected. Than animals were then boosted and bled every three weeks.

The ZSMF-7 peptide-specific seras were characterized by an ELISA titer check using 1 µg/ml of the peptide used to make the antibody (SEQ ID NOs: 35 and 36) as an antibody target. All 5 mouse seras to huzsmf7-2 and huzsmf7-3 have titer to their specific peptides at a dilution of 1 x 10^5. A single rabbit sera to huzsmf7-2 had titer to its specific peptide at a dilution of 1 x 10^5 and to recombinant full-length protein at a dilution of 1 x 10^5.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit

and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

We claim:

1. An isolated semaphorin polypeptide comprising a sequence of amino acid residues that is at least 80% identical in amino acid sequence to residues 69 through 541 of SEQ ID NO:2, said polypeptide comprising cysteine residues at positions corresponding to residues 126, 143, 152, 266, 291, 335, 366, 493, 500, 503, 511, 518 and 541 of SEQ ID NO:2.

2. An isolated semaphorin polypeptide according to claim 1, wherein said sequence of amino acid residues is at least 90% identical.

3. An isolated semaphorin polypeptide according to claim 1, further comprising an Ig-like domain.

4. An isolated semaphorin polypeptide according to claim 3, wherein said Ig-like domain comprises a sequence of amino acids from residue 561-620 of SEQ ID NO:2.

5. An isolated semaphorin polypeptide according to claim 1, wherein said polypeptide comprises residues 45-666 of SEQ ID NO:2.

6. An isolated semaphorin polypeptide according to claim 1, wherein said sequence of amino acid residues comprises residues 1-666 of SEQ ID NO:2.

7. An isolated semaphorin polypeptide according to claim 1, wherein said sequence of amino acid residues is from 473-624 amino acid residues.

8. An isolated semaphorin polypeptide selected from the group consisting of:

a) a polypeptide comprising a sequence of amino acid residues from amino acid residue 45 to residue 666 of SEQ ID NO:2;

b) a polypeptide comprising a sequence of amino acid residues from amino acid residue 69 to residue 666 of SEQ ID NO:2;

c) a polypeptide comprising a sequence of amino acid residues from amino acid residue 69 to residue 541 of SEQ ID NO:2 and

d) a polypeptide comprising a sequence of amino acid residues from amino acid residue 1 to residue 666 of SEQ ID NO:2.

9. An isolated semaphorin polypeptide according to claim 1, wherein any difference between said amino acid sequence of said isolated polypeptide and said corresponding amino acid sequence of SEQ ID NO:2 is due to a conservative amino acid substitution.

10. An isolated semaphorin polypeptide according to claim 1, covalently linked to a moiety selected from the group consisting of affinity tags, toxins, radionucleotides, enzymes and fluorophores.

11. An isolated semaphorin polypeptide according to claim 10, wherein said moiety is an affinity tag selected from the group consisting of polyhistidine, FLAG, Glu-Glu, glutathione S transferase and an immunoglobulin heavy chain constant region.

12. An isolated semaphorin polypeptide according to claim 11 further comprising a proteolytic cleavage site between said sequence of amino acid residues and said affinity tag.

13. An expression vector comprising the following operably linked elements:

a transcription promoter;

a DNA segment encoding a semaphorin polypeptide according to claim 1; and

a transcriptional terminator.

14. An expression vector according to claim 13 further comprising a secretory signal sequence operably linked to said DNA segment.

15. An expression vector according the claim 14, wherein said secretory signal sequence encodes residues 1-44 of SEQ ID NO:2.

16. An expression vector according to claim 13, wherein said sequence of amino acid residues is at least 90% identical.

17. An expression vector according to claim 13, wherein said DNA segment encodes a semaphorin polypeptide comprising an Ig-like domain.

18. An expression vector according to claim 17, wherein said Ig-like domain comprises a sequence of amino acids from residue 561-620 of SEQ ID NO:2.

19. An expression vector according to claim 13, wherein said sequence of amino acid residues comprises residues 45-666 of SEQ ID NO:2.

20. An expression vector according to claim 13, wherein said DNA segment encodes a semaphorin polypeptide covalently linked to an affinity tag selected from the group consisting of polyhistidine, FLAG, Glu-Glu, glutathione S transferase and an immunoglobulin heavy chain constant region.

21. A cultured cell into which has been introduced an expression vector according to claim 13, wherein said cell expresses the polypeptide encoded by the DNA segment.

22. A method of producing a semaphorin protein comprising:

culturing a cell into which has been introduced an expression vector according to claim 13, whereby said cell expresses said semaphorin protein encoded by said DNA segment; and

recovering said expressed semaphorin protein.

23. A pharmaceutical composition comprising a polypeptide according to claim 1, in combination with a pharmaceutically acceptable vehicle.

24. An antibody or antibody fragment that specifically binds to an epitope of a semaphorin polypeptide according to claim 1.

25. An antibody according to claim 24, wherein said antibody is selected from the group consisting of:

a) polyclonal antibody;
b) murine monoclonal antibody;
c) humanized antibody derived from b); and
d) human monoclonal antibody.

26. An antibody fragment according to claim 25, wherein said antibody fragment is selected from the group consisting of F(ab'), F(ab), Fab', Fab, Fv, scFv, and minimal recognition unit.

27. A binding protein that specifically binds to an epitope of a semaphorin polypeptide according to claim 1.

28. An anti-idiotype antibody that specifically binds to said antibody of claim 24.

29. An isolated polynucleotide encoding a semaphorin polypeptide according to claim 1.

30. An isolated polynucleotide according to claim 29, wherein said sequence of amino acid residues is at least 90% identical.

31. An isolated polynucleotide according to claim 29, wherein said semaphorin polypeptide comprises an Ig-like domain.

32. An isolated polynucleotide according to claim 31, wherein said Ig-like domain comprises a sequence of amino acids from residue 561-620 of SEQ ID NO:2.

33. An isolated polynucleotide according to claim 29, wherein said sequence of amino acid residues comprises residues 45-666 of SEQ ID NO:2.

34. An isolated polynucleotide according to claim 29, wherein said sequence of amino acid residues comprises residues 1-666 of SEQ ID NO:2.

35. An isolated polynucleotide according to claim 29 comprising nucleotide 1 to nucleotide 1998 of SEQ ID NO:5.

36. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 152 to nucleotide 2017 of SEQ ID NO:1;

b) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 244 to nucleotide 2017 of SEQ ID NO:1;

c) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 244 to nucleotide 1640 of SEQ ID NO:1;

d) a polynucleotide sequence consisting of the polynucleotide sequence from nucleotide 20 to nucleotide 2017 of SEQ ID NO:1; and

e) a complementary polynucleotide sequence of a, b, c or d.

37. A method for detecting a genetic abnormality in a patient, comprising:

obtaining a genetic sample from a patient;

incubating the genetic sample with a polynucleotide comprising at least 14 contiguous nucleotides of SEQ ID NO:1 or the complement of SEQ ID NO:1, under conditions wherein said polynucleotide will hybridize to complementary polynucleotide sequence, to produce a first reaction product;

comparing said first reaction product to a control reaction product, wherein a difference between said first reaction product and said control reaction product is indicative of a genetic abnormality in the patient.

```
MsemF     -------------------------------------------MAPHWAV
MsemE     ---------------------------------------------MAFRA
MsemC     --------------------------------------------------
ZSMF7     ------------MTPPPPGRAAPSAPRAR------------VPGPPARLG
AHU18243  MAYLNATVSKPVISLLSLSKKVLKFEHCGGEGQCLGLITEFVIHPAAMGT
MsemD     -------------------------------------------MGWFTGI
MsemA     -------------------------------------------MGRAEAA
MsemB     ------------------------------------MALPSLGQDSWSLL

MsemF     WLLAAGLWGLGIGAEMWWNL-VPRKTVSSGELVT-----VVRRFSQTGI-
MsemE     ICVLVGVFICSICVRGSSQP-QARVYLTFDELRETKT---SEYFSLSHQQ
MsemC     ------------------------------EER------LIRKFEAENI-
ZSMF7     LPLRLRLLLLLWAAAASAQG-HLRSGPRIFAVWKGHVGQDRVDFGQT---
AHU18243  LCVSIRLLMILSAITAAKSRFIDKPR-LIVNLTDGFG--QHRFFGPQ---
MsemD     ACLFWGVLLTARANYANGKNNVPRLKLSYKEMLESNN--VITFNGLANS-
MsemA     VMIP-GLALLWVAGLGDTAPNLPRLRLSFQELQARH---GVRTFRLERT-
MsemB     RVFFFQLFLLPSLPPASGTGGQGPMPRVKYHAGDGHR--ALSFFQQKGL-

MsemF     QDFLTLTLTEHSGLLYVGAREALFAFSVEALEL----QGAISWEAPAEKK
MsemE     LDYRILLMDEDQDRIYVGSKDHILSLNINNISQ---EPLSVFWPASTIKV
MsemC     SNYTALLLSQDGKTLYVGAREALFALNSNLSFLPGGEYQELLWSADADRK
ZSMF7     EPHTVLFHEPGSSSVWVGGRGKVYLFDFPEG------KNASVRTVNIGST
AHU18243  EPHTVLFHSLNSSDVYVGGNNTIYLFDFAHS------SNASTALINITST
MsemD     SSYHTFLLDEERSRLYVGAKDHIFSFNLVNIK----DFQKIVWPVSYTRR
MsemA     CCYEALLVDEERGRLFVGAENHVASLSLDNISK---RAKKLAWPAPVEWR
MsemB     RDFDTLLLSDDGNTLYVGARETVLALNIQNPGIP-RLKNMIPWPASERKK
            .  :        ::**..  :  :.

MsemF     IECTQKGKSNQTECFNFIRFLQPYNSSHLYVCGTYAFQPKCTYINMLTFT
MsemE     EECKMAGKDPTHGCGNFVRVIQTFNRTHLYVCGSGAFSPVCTYLNRGRRS
MsemC     QQCSFKGKDPKRDCQNYIKILLPLNSSHLLTCGTAAFSPLCAYIHIASFT
ZSMF7     KGSCLDKRD----CENYITLLERR-SEGLLACGTNARHPSCWNLVNGTVV
AHU18243  HNTHRLSST----CENFITLLHNQ-TDGLLACGTNSQKPSCWLINNLTTQ
MsemD     DECKWAGKDILKECANFIKVLEAYNQTHLYACGTGAFHPICTYIEVGHHP
MsemA     EECNWAGKDIGTECMNFVRLLHAYNHTHLLACRTGAFHPTCALWRWATAG
MsemB     TECAFKKKSNETQCFNFIRVLVSYNATHLYACGTFAFSPACTFIELQDSL
                       * *:: .:       * .* : :  * *

MsemF     LDRAEF------EDGKGKCPYDPAKGHTGLLVDGELYSATLNNFLGTEPV
MsemE     EDQVF-MIDSKCESGKGRCSFNPNVNTVSVMINEELFSGMYIDFMGTDAA
MsemC     LAQDEAGNVI-LEDGKGHCPFDPNFKSTALVVDGELYTGTVSSFQGNDPA
ZSMF7     PLGEM----------RGYAPFSPDENSLVLFEGDEVYSTIRKQEYNGKIP
AHU18243  FLGPK----------LGLAPFSPSSGNLVLFDQNDTYSTINLYKSLSGSH
MsemD     EDNIFKLQDSHFENGRGKSPYDPKLLTASLLIDGELYSGTAADFMGRDFA
MsemA     GTHAS-TGPEKLEDGKGKTPYDPRHRPPSVLVGEELYSGVTADLMGRDFT
MsemB     LLPILIDK---VMDGKGQSPLTLFTSTQAVLVDGMLYSGTMNNFLGSEPI
                          *  .         :.     ::

MsemF     ILRYMGTHHSIKTEYL-AFWLNEPHFVGSAFVPESVGSFTGDDDKIYFFF
MsemE     IFRSLTKRMQLRTDQHNSKWLSEPMFVDAHVIPDGTDP---NDAKVYFFF
MsemC     ISRSQ-SSRPTKTESS-LNWLQDPAFVASATSPESLGSPIGDDDKIYFFF
ZSMF7     RFRRIRGESELYTSDT---VMQNPQFIKATIVHQDQA----YDDKIYYFF
AHU18243  KFRRIAGQVELYTSDT---AMHRPQFVQATAVHKNES----YDDKIYFFF
MsemD     IFRTLGDHHPIRTEQHDSRWLNDPRFISAHLIPESDNP---EDDKVYFFF
MsemA     IFRSLGQNPSLRTEPHDSRWLNEPKFVKVFWIPESENP---DDDKIYFFF
MsemB     LMRTLGSHPVLKTDIF-LRWLHADASF----VAAIPS-----TQVVYFFF
            *         *.      :     .                  :*:**
```

Figure 1a

```
MsemF     SERAVEYD-CYSEQVVARVARVCKGDMGGARTL-QKKWTTFLKARLVCSA
MsemE     KERLTDNN-RSTKQIHSMIARICPNDTGGQRSL-VNKWTTFLKARLVCSV
MsemC     SETGQEFE-FFENTIVSRVARVCKGDEGGERVL-QQRWTSFLKAQLLCSR
ZSMF7     REDNPDKN-PEAPLNVSRVAQLCRGDQGGESSLSVSKWNTFLKAMLVCSD
AHU18243  QENSHSDF-KQFPHTVPRVGQVCSSDQGGESSLSVYKWTTFLKARLACVD
MsemD     RENAIGGE-HSGKATHARIGQICKNDFGGHRSL-VNKWTTFLKARLICSV
MsemA     RESAVEAAPAMGRMSVSRVGQICRNDLGGQRSL-VNKWTTFLKARLVCSV
MsemB     EETASEFD-FFEELYISRVAQVCKNDVGGEKLL-QKKWTTFLKAQLLCAQ
           *              . :.::* .* **   *   :*.:**** * *

MsemF     PDWKV---YFNQLKAVHTLR--GASWHNTTFFGVFQARWGD--MDLSAVC
MsemE     TDEDGPETHFDELEDVFLLE--TDNPRTTLVYGIFTTSSSV--FKGSAVC
MsemC     PDDGFP---FNVLQDVFTLNPNPQDWRKTLSIGVFTSQWHRGTTEGSAIC
ZSMF7     AATNK---NFNRLQDVFLLPDPSGQWRDTRVYGVFSNPWN-----YSAVC
AHU18243  YDTGR---IYNELQDIFIWQAPENSWEETLIYGLFLSPWN-----FSAVC
MsemD     PGPNGIDTHFDELQDVFLMN--SKDPKNPIVYGVFTTSSNI--FKGSAVC
MsemA     PGVEG-DTHFDQLQDVFLLS--SRDRQTPLLYAVFSTSSGV--FQGSAVC
MsemB     PGQLP----FNIIRHAVLLP--ADSPSVSRIYAVFTSQWQVGGTRSSAVC
                   :: :.          .   .   .:*           **:*

MsemF     EYQLEQIQQVFEGPYKEYSEQAQKWARYTDPVPSPRPGSCINNWHRDNGY
MsemE     VYHLSDIQTVFNGPFAHKEGPNHQLISYQGRIPYPRPGTCPGGAFTP-NM
MsemC     VFTMNDVQKAFDGLYKKVNRETQQWYTETHQVPTPRPGACITNSARERKI
ZSMF7     VYSLGDIDKVFR---TS------SLKGYHSSLPNPRPGKCLPDQQP----
AHU18243  VFTVKDIDHVFK---TS------KLKNYHHKLPTPRPGQCMKNHQH----
MsemD     MYSMSDVRRVFLGPYAHRDGPNYQWVPYQGRVPYPRPGTCPSKTFG--GF
MsemA     VYSMNDVRRAFLGPLPHKEGPTHQWVSYQGRVPYPRPGMCPSKTFG--TF
MsemB     AFSLTDIERVFKGKYKELNKETSRWTTYRGSEVSPRPGSCSMGPSS----
           : : ::  .*                       **** *

MsemF     TSSLELPDNTLNFIKKHPLMEDQVKPRL-GRPLLVKKNTNFTH--VVADR
MsemE     RTTKDFPDDVVTFIRNHPLMYNSISPIH-RRPLIVRIGTDYKYTKIAVDR
MsemC     NSSLQLPDRVLNFLKDHFLMDGQVRSRL----LLLQPRARYQR--VAVHR
ZSMF7     -----IPTETFQVADRHPEVAQRVEPMGPLKTPLFHSKYHYQK--VAVHR
AHU18243  -----VPTETFQVADRYPEVADPVYQKNNAMFPIIQSKYIYTK--LLVYR
MsemD     DSTKDLPDDVITFGRSHPAMYNPVFPIN-NRPIMIKTDVNYQFTQIVVDR
MsemA     SSTKDFPDDVIQFGRNHPLMYNPVLPMG-GRPLFLQVGAGYTFTQIAADR
MsemB     -------DKALTFMKDHFLMDEHVVGT----PLLVKSGVEYTR--LAVES
                   .. .   :  :   :         :.:    :     : .

MsemF     VPGLDGATYTVLFIGTGDGWLLKAVS------LGPWIHMVEELQVFDQ-E
MsemE     VNAAD-GRYHVLFLGTDRGTVQKVVVLPTNSSASG-ELILEELEVFKNHV
MsemC     VPGLH-STYDVLFLGTGDGRLHKAVT------LSSRVHIIEELQIFPQGQ
ZSMF7     MQASHGETFHVLYLTTDRGTIHKVVEP--GEQEHSFAFNIMEIQPFRRAA
AHU18243  VEYGGVFWATIFYLTTIKGTIHIYVRY--EDSNSTTALNILEINPFQKPA
MsemD     VDAED-GQYDVMFIGTDVGTVLKVVSVPKETWHDLEEVLLEEMTVFREPT
MsemA     VAAAD-GHYDVLFIGTDVGTVLKVISVPKGRRPNSEGLLLEELQVFEDSA
MsemB     ARGLDGSSHVVMYLGTSTGPLHKAVVP-----QDSSAYLVEEIQLSPDSE
                    :::: *  * :   :              : *:

MsemF     PVESLVLSQSKKVLFAGSRSQLVQLSLADCTKY-RFCVDCVLARDPYCAW
MsemE     PITTMEISSKKQQLYVSSNEGVSQVSLHRCHIYGTACADCCLARDPYCAW
MsemC     PVQNLLLDSHGGLLYASSHSGVVQVPVANCSLY-PTCGDCLLARDPYCAW
ZSMF7     AIQTMSLDAERRKLYVSSQWEVSQVPLDLCEVYGGGCHGCLMSRDPYCGW
AHU18243  PIQNILLDNTNLKLYVNSEWEVSEVPLDLCSVYGNDCFSCFMSRDPLCTW
MsemD     TISAMELSTKQQQLYIGSTAGVAQLPLHRCDIYGKACAECCLARDPYCAW
MsemA     AITSMQISSKRQQLYVASRAAVAQIALHRCTALGRACAECCLARDPYCAW
MsemB     PVRNLQLAPAQGAVFAGFSGGIWRVPRANCSVY-ESCVDCVLARDPHCAW
          .:  : :      ::      : .:.   *      *  * ::*** * *
```

Figure 1b

```
MsemF       NVNTSRC---VATTSGRSGSFLVQHVANLDTSKMCN------QYGIKKVR
MsemE       DGHS--C---SRFYPTGKRRSRRQDVRHGNPLTQCRG---F-NLKAYRNA
MsemC       TGSA--CRLASLYQPDLASRPWTQDIEGASVKELCKN-SSY-KARFLVPG
ZSMF7       DQGR--C--ISIYSSE---RSVLQSINPAEPHKECP------NPKPDK--
AHU18243    YNNT--C---SFKQRV---SVETGGPANRTLSEMCG--------DHYAPT
MsemD       DGSS--C---SRYFPTAKRRTRRQDIRNGDPLTHCSDLEDH-DNHHGPSL
MsemA       DGSA--C---TRFQPTAKRRFRRQDIRNGDPSTLCS----G-DSSHSVLL
MsemB       DPESRLC---SLLSGS--TKPWKQDMERGNPEWVCTRGPMARSPRRQSPP
                  *                            *

MsemF       SIPKNITVVSGTDLVLPCHLSSNLAHAHWTFGS-QDLP--AEQP-GSFLY
MsemE       AEIVQYGVR-NNSTFLECAPKSPQASIKWLLQKDKDRR--KEGKLNERII
MsemC       KPCKQVQIQPNTVNTLACPLLSNLATRLWVHNG-APVN----ASASCRVL
ZSMF7       APLQKVSLAPNSRYYLSCPMESRHATYSWRHKENVEQS------CEPGHQ
AHU18243    VVKHQVSIPLLSNSYLSCPAVSNHADYFWTKDGFTEKR------CHVKTH
MsemD       EERIIYGVE-NSSTFLECSPKSQRALVYWQFQR-RNRRSKREIRMGDHII
MsemA       EKKVL-GVE-SGSAFLECEPRSLQAHVQWTFQG-AGEAAHTQVLAEERVE
MsemB       QLIKEVLTVPNSILELRCPHLSALASYHWSHGR-AKIS-----E-ASATV
                             * *   *  *   *

MsemF       DTGLQALVVMAAQSRHSGPYRCYSEEQGTRLAAESYLVAVVAGS----SV
MsemE       AT-SQGLLIRSVQDSDQGLYHCIATENS--FKQTIAKINFKVLD-----S
MsemC       PT---GDLLLVGSQQGLGVFQCWSIEEG--FQQLVASYCPEVME----EG
ZSMF7       SP-NCILFIENLTAQQYGHYFCEAQEGS--YFREAQHWQLLPED--GIMA
AHU18243    KN-DCILLIANSTTATNGTHVCNMKEDS--VTVKLLEVNVTLM-------
MsemD       RT-EQGLLLRSLQKKDSGNYLCHAVEHG--FMQTLLKVTLEVID-TEHLE
MsemA       RT-ARGLLLRGLRRQDSGVYLCVAVEQG--FSQPLRRLVLHVLS------
MsemB       YN---GSLLLLPQDGVGGLYQCVATENG--YSYPVVSYWVDSQDQPLALD
                  .:           * . *   *  .

MsemF       TLEARAPLENLGLVWLAVVALGAVC-LVLLLLVLSLRRRLREELEKGAKA
MsemE       EMVAVVTDKWSPWTWAGSVRALP----FHPKDILGAFS----HSEMQLIN
MsemC       VMDQKNQRDGTPVIINTSRVSAPAGGRDSWGADKSYWNEFLVMCTLFVFA
ZSMF7       EHLLGHACALAASLWLGVLPTLTLGLLVH---------------------
AHU18243    --------------------------------------------------
MsemD       ELLHKDDDGDGSKIKEMSSSMTPSQ-KVWYRDFMQLIN----HPNLNTMD
MsemA       ----AAQAERLARAEEAAAPAPPGP-KLWYRDFLQLVE----PGGGGGAN
MsemB       PELAGVPRERVQVPLTRVGGGASMAAQRSYWPHFLIVTVLLAIVLLGVLT
```

Figure 1c

SEQUENCE LISTING

<110> ZymoGenetics, Inc.
1201 Eastlake Avenue East
Seattle, Washington 98102
United States of America

<120> HUMAN SEMAPHORIN ZSMF-7

<130> 97-59

<150> 60/076.611
<151> 1998-03-03

<160> 36

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 3377
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (20)...(2017)

<400> 1
gaattcggca cgaggcggg atg acg cct cct ccg ccc gga cgt gcc gcc ccc 52
Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro
1 5 10

agc gca ccg cgc gcc cgc gtc cct ggc ccg ccg gct cgg ttg ggg ctt 100
Ser Ala Pro Arg Ala Arg Val Pro Gly Pro Pro Ala Arg Leu Gly Leu
15 20 25

ccg ctg cgg ctg cgg ctg ctg ctg ctg ctc tgg gcg gcc gcc gcc tcc 148
Pro Leu Arg Leu Arg Leu Leu Leu Leu Leu Trp Ala Ala Ala Ala Ser
30 35 40

gcc cag ggc cac cta agg agc gga ccc cgc atc ttc gcc gtc tgg aaa 196
Ala Gln Gly His Leu Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys

45 50 55

ggc cat gta ggg cag gac cgg gtg gac ttt ggc cag act gag ccg cac 244
Gly His Val Gly Gln Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His
60 65 70 75

acg gtg ctt ttc cac gag cca ggc agc tcc tct gtg tgg gtg gga gga 292
Thr Val Leu Phe His Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly
80 85 90

cgt ggc aag gtc tac ctc ttt gac ttc ccc gag ggc aag aac gca tct 340
Arg Gly Lys Val Tyr Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser
95 100 105

gtg cgc acg gtg aat atc ggc tcc aca aag ggg tcc tgt ctg gat aag 388
Val Arg Thr Val Asn Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys
110 115 120

cgg gac tgc gag aac tac atc act ctc ctg gag agg cgg agt gag ggg 436
Arg Asp Cys Glu Asn Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly
125 130 135

ctg ctg gcc tgt ggc acc aac gcc cgg cac ccc agc tgc tgg aac ctg 484
Leu Leu Ala Cys Gly Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu
140 145 150 155

gtg aat ggc act gtg gtg cca ctt ggc gag atg aga ggc tac gcc ccc 532
Val Asn Gly Thr Val Val Pro Leu Gly Glu Met Arg Gly Tyr Ala Pro
160 165 170

ttc agc ccg gac gag aac tcc ctg gtt ctg ttt gaa ggg gac gag gtg 580
Phe Ser Pro Asp Glu Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val
175 180 185

tat tcc acc atc cgg aag cag gaa tac aat ggg aag atc cct cgg ttc 628
Tyr Ser Thr Ile Arg Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe
190 195 200

cgc cgc atc cgg ggc gag agt gag ctg tac acc agt gat act gtc atg 676
Arg Arg Ile Arg Gly Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met
205 210 215

cag aac cca cag ttc atc aaa gcc acc atc gtg cac caa gac cag gct 724
Gln Asn Pro Gln Phe Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala
220 225 230 235

```
tac gat gac aag atc tac tac ttc ttc cga gag gac aat cct gac aag      772
Tyr Asp Asp Lys Ile Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys
                240                 245                 250

aat cct gag gct cct ctc aat gtg tcc cgt gtg gcc cag ttg tgc agg      820
Asn Pro Glu Ala Pro Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg
            255                 260                 265

ggg gac cag ggt ggg gaa agt tca ctg tca gtc tcc aag tgg aac act      868
Gly Asp Gln Gly Gly Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr
        270                 275                 280

ttt ctg aaa gcc atg ctg gta tgc agt gat gct gcc acc aac aag aac      916
Phe Leu Lys Ala Met Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn
    285                 290                 295

ttc aac agg ctg caa gac gtc ttc ctg ctc cct gac ccc agc ggc cag      964
Phe Asn Arg Leu Gln Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln
300                 305                 310                 315

tgg agg gac acc agg gtc tat ggt gtt ttc tcc aac ccc tgg aac tac     1012
Trp Arg Asp Thr Arg Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr
                320                 325                 330

tca gcc gtc tgt gtg tat tcc ctc ggt gac att gac aag gtc ttc cgt     1060
Ser Ala Val Cys Val Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg
            335                 340                 345

acc tcc tca ctc aag ggc tac cac tca agc ctt ccc aac ccg cgg cct     1108
Thr Ser Ser Leu Lys Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro
        350                 355                 360

ggc aag tgc ctc cca gac cag cag ccg ata ccc aca gag acc ttc cag     1156
Gly Lys Cys Leu Pro Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln
    365                 370                 375

gtg gct gac cgt cac cca gag gtg gcg cag agg gtg gag ccc atg ggg     1204
Val Ala Asp Arg His Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly
380                 385                 390                 395

cct ctg aag acg cca ttg ttc cac tct aaa tac cac tac cag aaa gtg     1252
Pro Leu Lys Thr Pro Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val
                400                 405                 410
```

```
gcc gtc cac cgc atg caa gcc agc cac ggg gag acc ttt cat gtg ctt      1300
Ala Val His Arg Met Gln Ala Ser His Gly Glu Thr Phe His Val Leu
            415                     420                     425

tac cta act aca gac agg ggc act atc cac aag gtg gtg gaa ccg ggg      1348
Tyr Leu Thr Thr Asp Arg Gly Thr Ile His Lys Val Val Glu Pro Gly
        430                     435                     440

gag cag gag cac agc ttc gcc ttc aac atc atg gag atc cag ccc ttc      1396
Glu Gln Glu His Ser Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe
    445                     450                     455

cgc cgc gcg gct gcc atc cag acc atg tcg ctg gat gct gag cgg agg      1444
Arg Arg Ala Ala Ala Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg
460                     465                     470                     475

aag ctg tat gtg agc tcc cag tgg gag gtg agc cag gtg ccc ctg gac      1492
Lys Leu Tyr Val Ser Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp
                480                     485                     490

ctg tgt gag gtc tat ggc ggg ggc tgc cac ggt tgc ctc atg tcc cga      1540
Leu Cys Glu Val Tyr Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg
            495                     500                     505

gac ccc tac tgc ggc tgg gac caa ggc cgc tgc atc tct atc tac agc      1588
Asp Pro Tyr Cys Gly Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser
        510                     515                     520

tcc gaa cgg tca gtg ctg caa tcc att aat cca gcc gag cca cac aag      1636
Ser Glu Arg Ser Val Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys
    525                     530                     535

gag tgt ccc aac ccc aaa cca gac aag gcc cca ctg cag aag gtt tcc      1684
Glu Cys Pro Asn Pro Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser
540                     545                     550                     555

ctg gcc cca aac tct cgc tac tac ctg agc tgc ccc atg gaa tcc cgc      1732
Leu Ala Pro Asn Ser Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg
                560                     565                     570

cac gcc acc tac tca tgg cgc cac aag gag aac gtg gag cag agc tgc      1780
His Ala Thr Tyr Ser Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys
            575                     580                     585

gaa cct ggt cac cag agc ccc aac tgc atc ctg ttc atc gag aac ctc      1828
```

Glu Pro Gly His Gln Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu
590 595 600

acg gcg cag cag tac ggc cac tac ttc tgc gag gcc cag gag ggc tcc 1876
Thr Ala Gln Gln Tyr Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser
605 610 615

tac ttc cgc gag gct cag cac tgg cag ctg ctg ccc gag gac ggc atc 1924
Tyr Phe Arg Glu Ala Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile
620 625 630 635

atg gcc gag cac ctg ctg ggt cat gcc tgt gcc ctg gcc gcc tcc ctc 1972
Met Ala Glu His Leu Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu
640 645 650

tgg ctg ggg gtg ctg ccc aca ctc act ctt ggc ttg ctg gtc cac 2017
Trp Leu Gly Val Leu Pro Thr Leu Thr Leu Gly Leu Leu Val His
655 660 665

tagggcctcc cgaggctggg catgcctcag gcttctgcag cccagggcac tagaacgtct 2077
cacactcaga gccggctggc ccgggagctc cttgcctgcc acttcttcca ggggacagaa 2137
taacccagtg gaggatgcca ggcctggaga cgtccagccg caggcggctg ctgggcccca 2197
ggtggcgcac ggatggtgag gggctgagaa tgagggcacc gactgtgaag ctggggcatc 2257
gatgacccaa gactttatct tctggaaaat atttttcaga ctcctcaaac ttgactaaat 2317
gcagcgatgc tcccagccca agagcccatg ggtcggggag tgggtttgga taggagagct 2377
gggactccat ctcgaccctg gggctgaggc ctgagtcctt ctggactctt ggtacccaca 2437
ttgcctcctt cccctccctc tctcatggct gggtggctgg tgttcctgaa gacccagggc 2497
taccctctgt ccagccctgt cctctgcagc tccctctctg gtcctgggtc ccacaggaca 2557
gccgccttgc atgtttattg aaggatgttt gctttccgga cggaaggacg gaaaaagctc 2617
tatttttatg ttaggcttat ttcatgtata gctacttccg actgcatctg tatgaaaata 2677
ccaaaactac atgcgggggg gtgggtggga aagggagggg ctgggaaggg atgggttggg 2737
gagcgggggt gatcccagtc tgaggctccc ggggatgaga taagagtctg gagacgggca 2797
tgggttcttg gagagtggca tgagctggct ctgccctggg agcccggtct gagggggacg 2857
ttgttggagc ccctagtgtt gggggtggtt atgggagggg gtggggtgag ggaaacggga 2917
gaatgaagga gaaaactgag ccctagtttc accgtgttca tttggaagga cgagccgggt 2977
cctcaggggg aggttccagg actctgccct tggcgttgag ggttgggggg cggggggcct 3037
cctcccttcc tctcagcccc cttccccagg ggctgtgctt ccatgctcct agcctcccac 3097
cttcgctcag gacatgttat aacttaggct aaactgtgaa aattccggtg ggatggcct 3157
gggccgagct ctccaggcag gcggccctgc ccccagccct gtccatccat ttcagggggg 3217
agctgggccc ttctccggct gtgtctggcc acccagggca gtggctgggg ccagtggcct 3277
tccagctttg gcccctgcac ctcttctcaa tgcactttaa taatgtaaca tattactaat 3337
aaacaagcta tttatttaaa aaaaaaaaaa aaaagagctc 3377

<210> 2
<211> 666

<212> PRT
<213> Homo sapiens

<400> 2
Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala
1   5   10   15
Arg Val Pro Gly Pro Pro Ala Arg Leu Gly Leu Pro Leu Arg Leu Arg
            20   25   30
Leu Leu Leu Leu Leu Trp Ala Ala Ala Ala Ser Ala Gln Gly His Leu
        35   40   45
Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys Gly His Val Gly Gln
    50   55   60
Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu Phe His
65   70   75   80
Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr
                85   90   95
Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn
            100   105   110
Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu Asn
        115   120   125
Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys Gly
    130   135   140
Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Gly Thr Val
145   150   155   160
Val Pro Leu Gly Glu Met Arg Gly Tyr Ala Pro Phe Ser Pro Asp Glu
                165   170   175
Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg
            180   185   190
Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly
        195   200   205
Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe
    210   215   220
Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile
225   230   235   240
Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro
                245   250   255
Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly
            260   265   270
Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met
        275   280   285
Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn Phe Asn Arg Leu Gln
    290   295   300
Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg
305   310   315   320
Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val

```
                325                     330                     335
Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys
            340                     345                     350
Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro
        355                     360                     365
Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His
    370                     375                     380
Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro
385                     390                     395                     400
Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met
                405                     410                     415
Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp
            420                     425                     430
Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser
        435                     440                     445
Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala
    450                     455                     460
Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser
465                     470                     475                     480
Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr
                485                     490                     495
Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly
            500                     505                     510
Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val
        515                     520                     525
Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro
    530                     535                     540
Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser
545                     550                     555                     560
Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser
                565                     570                     575
Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln
            580                     585                     590
Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr
        595                     600                     605
Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala
    610                     615                     620
Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu
625                     630                     635                     640
Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu
                645                     650                     655
Pro Thr Leu Thr Leu Gly Leu Leu Val His
            660                     665
```

<210> 3

<211> 390
<212> PRT
<213> Mus musculus

<400> 3

```
Met Asn Lys Leu Leu Cys Cys Ala Leu Val Phe Leu Asp Ile Ser Ile
 1               5                  10                  15
Lys Trp Thr Thr Gln Glu Thr Phe Pro Pro Lys Tyr Leu His Tyr Asp
            20                  25                  30
Glu Glu Thr Ser His Gln Leu Leu Cys Asp Lys Cys Pro Pro Gly Thr
        35                  40                  45
Tyr Leu Lys Gln His Cys Thr Ala Lys Trp Lys Thr Val Cys Ala Pro
    50                  55                  60
Cys Pro Asp His Tyr Tyr Thr Asp Ser Trp His Thr Ser Asp Glu Cys
65                  70                  75                  80
Leu Tyr Cys Ser Pro Val Cys Lys Glu Leu Gln Tyr Val Lys Gln Glu
                85                  90                  95
Cys Asn Arg Thr His Asn Arg Val Cys Glu Cys Lys Glu Gly Arg Tyr
            100                 105                 110
Leu Glu Ile Glu Phe Cys Leu Lys His Arg Ser Cys Pro Pro Gly Phe
        115                 120                 125
Gly Val Val Gln Ala Gly Thr Pro Glu Arg Asn Thr Val Cys Lys Arg
    130                 135                 140
Cys Pro Asp Gly Phe Phe Ser Asn Glu Thr Ser Ser Lys Ala Pro Cys
145                 150                 155                 160
Arg Lys His Thr Asn Cys Ser Val Phe Gly Leu Leu Leu Thr Gln Lys
                165                 170                 175
Gly Asn Ala Thr His Asp Asn Ile Cys Ser Gly Asn Ser Glu Ser Thr
            180                 185                 190
Gln Lys Cys Gly Ile Asp Val Thr Leu Cys Glu Glu Ala Phe Phe Arg
        195                 200                 205
Phe Ala Val Pro Thr Lys Phe Thr Pro Asn Trp Leu Ser Val Leu Val
    210                 215                 220
Asp Asn Leu Pro Gly Thr Lys Val Asn Ala Glu Ser Val Glu Arg Ile
225                 230                 235                 240
Lys Arg Gln His Ser Ser Gln Glu Gln Thr Phe Gln Leu Leu Lys Leu
                245                 250                 255
Trp Lys His Gln Asn Lys Asp Gln Asp Ile Val Lys Lys Ile Ile Gln
            260                 265                 270
Asp Ile Asp Leu Cys Glu Asn Ser Val Gln Arg His Ile Gly His Ala
        275                 280                 285
Asn Leu Thr Phe Glu Gln Leu Arg Ser Leu Met Glu Ser Leu Pro Gly
    290                 295                 300
Lys Lys Val Gly Ala Glu Asp Ile Glu Lys Thr Ile Lys Ala Cys Lys
305                 310                 315                 320
```

```
Pro Ser Asp Gln Ile Leu Lys Leu Leu Ser Leu Trp Arg Ile Lys Asn
                325                 330                 335
Gly Asp Gln Asp Thr Leu Lys Gly Leu Met His Ala Leu Lys His Ser
            340                 345                 350
Lys Thr Tyr His Phe Pro Thr Asn Cys His Ser Glu Ser Lys Glu Asp
        355                 360                 365
His Gln Val Pro Ser Gln Leu His Asn Val Gln Ile Val Ser Glu Val
    370                 375                 380
Ile Phe Arg Asn Asp Arg
385                 390
```

<210> 4
<211> 233
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide probe for Northern Blots

<400> 4

```
gtctggaaag gccatgtagg gcaggaccgg gtggactttg gccagactga gccgcacacg     60
gtgcttttcc acgagccagg cagctcctct gtgtgggtgg gaggacgtgg caaggtctac    120
ctctttgact tccccgaggg caagaacgca tctgtgcgca cggtgaatat cggctccaca    180
aaggggtcct gtctggataa gcgggactgc gagaactaca tcactctcct gga           233
```

<210> 5
<211> 1998
<212> DNA
<213> Artificial Sequence

<220>
<223> Degenerate oligonucleotide sequence encoding the zsmf7 polypeptide of SEQ ID NO:2

<221> variation
<222> (1)...(1998)
<223> Each N is independently any nucleotide.

<400> 5

```
atgacnccnc cnccnccngg nmgngcngcn ccnwsngcnc cnmgngcnmg ngtnccnggn     60
ccnccngcnm gnytnggnyt nccnytnmgn ytnmgnytny tnytnytnyt ntgggcngcn    120
gcngcnwsng cncarggnca yytnmgnwsn ggnccnmgna thttygcngt ntggaarggn    180
caygtnggnc argaymgngt ngayttyggn caracngarc cncayacngt nytnttycay    240
garccnggnw snwsnwsngt ntgggtnggn ggnmgnggna argtntayyt nttygaytty    300
```

ccngarggna araaygcnws ngtnmgnacn gtnaayathg gnwsnacnaa rggnwsntgy 360
ytngayaarm gngaytgyga raaytayath acnytnytng armgnmgnws ngarggnytn 420
ytngcntgyg gnacnaaygc nmgncayccn wsntgytgga ayytngtnaa yggnacngtn 480
gtnccnytng gngaratgmg nggntaygcn ccnttywsnc cngaygaraa ywsnytngtn 540
ytnttygarg gngaygargt ntaywsnacn athmgnaarc argartayaa yggnaarath 600
ccnmgnttym gnmgnathmg nggngarwsn garytntaya cnwsngayac ngtnatgcar 660
aayccncart tyathaargc nacnathgtn caycargayc argcntayga ygayaarath 720
taytayttyt tymgngarga yaayccngay aaraayccng argcnccnyt naaygtnwsn 780
mgngtngcnc arytntgymg nggngaycar ggnggngarw snwsnytnws ngtnwsnaar 840
tggaayacnt tyytnaargc natgytngtn tgywsngayg cngcnacnaa yaaraaytty 900
aaymgnytnc argaygtntt yytnytnccn gayccnwsng gncartggmg ngayacnmgn 960
gtntayggng tnttywsnaa yccntggaay taywsngcng tntgygtnta ywsnytnggn 1020
gayathgaya argtnttymg nacnwsnwsn ytnaarggnt aycaywsnws nytnccnaay 1080
ccnmgnccng gnaartgyyt nccngaycar carccnathc cnacngarac nttycargtn 1140
gcngaymgnc ayccngargt ngcncarmgn gtngarccna tgggnccnyt naaracnccn 1200
ytnttycayw snaartayca ytaycaraar gtngcngtnc aymgnatgca rgcnwsncay 1260
ggngaracnt tycaygtnyt ntayytnacn acngaymgng gnacnathca yaargtngtn 1320
garccnggng arcargarca ywsnttygcn ttyaayatha tggarathca rccnttymgn 1380
mgngcngcng cnathcarac natgwsnytn gaygcngarm gnmgnaaryt ntaygtnwsn 1440
wsncartggg argtnwsnca rgtnccnytn gayytntgyg argtntaygg nggnggntgy 1500
cayggntgyy tnatgwsnmg ngayccntay tgyggntggg aycarggnmg ntgyathwsn 1560
athtaywsnw sngarmgnws ngtnytncar wsnathaayc cngcngarcc ncayaargar 1620
tgyccnaayc cnaarccnga yaargcnccn ytncaraarg tnwsnytngc nccnaaywsn 1680
mgntaytayy tnwsntgycc natggarwsn mgncaygcna cntaywsntg gmgncayaar 1740
garaaygtng arcarwsntg ygarccnggn caycarwsnc cnaaytgyat hytnttyath 1800
garaayytna cngcncarca rtayggncay tayttytgyg argcncarga rggnwsntay 1860
ttymgngarg cncarcaytg gcarytnytn ccngargayg gnathatggc ngarcayytn 1920
ytnggncayg cntgygcnyt ngcngcnwsn ytntggytng gngtnytncc nacnytnacn 1980
ytnggnytny tngtncay 1998

<210> 6
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16086

<400> 6
aggaccgggt ggactttg 18

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC 16085

<400> 7
tcggggaagt caaagagg 18

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC694

<400> 8
taatacgact cactataggg 20

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16820

<400> 9
acacctcgtc cccttcaaac 20

<210> 10
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16087

<400> 10
aagcgggact gcgagaac 18

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16818

<400> 11
gttgggaagg cttgagtggt 20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide 15394

<400> 12
ctggagaggc ggagtgaggg 20

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16819

<400> 13
catgatgttg aaggcgaagc 20

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16460

<400> 14
tgatgctgcc accaacaaga 20

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16548

<400> 15
aagacgccat tgttccactc 20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16807

<400> 16
tgggcctcgc agaagtagtg 20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16806

<400> 17
attctcagcc cctcaccatc 20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16667

<400> 18
atggggccac aaggagaacg 20

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Oligonucleotide ZC16729

<400> 19
ccgtccggaa agcaaacatc 20

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16728

<400> 20
tccttgcctg ccacttcttc 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16666

<400> 21
cggctcgtcc ttccaaatga 20

<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC2681

<400> 22
gaataagagt atagaaga 18

<210> 23
<211> 0
<212> PRT
<213> Mus musculus

<400> 23
Met Ala Pro His Trp Ala Val Trp Leu Leu Ala Ala Gly Leu Trp Gly

Leu Gly Ile Gly Ala Glu Met Trp Trp Asn Leu Val Pro Arg Lys Thr

Val Ser Ser Gly Glu Leu Val Thr Val Val Arg Arg Phe Ser Gln Thr

Gly Ile Gln Asp Phe Leu Thr Leu Thr Leu Thr Glu His Ser Gly Leu

Leu Tyr Val Gly Ala Arg Glu Ala Leu Phe Ala Phe Ser Val Glu Ala

Leu Glu Leu Gln Gly Ala Ile Ser Trp Glu Ala Pro Ala Glu Lys Lys

Ile Glu Cys Thr Gln Lys Gly Lys Ser Asn Gln Thr Glu Cys Phe Asn

Phe Ile Arg Phe Leu Gln Pro Tyr Asn Ser Ser His Leu Tyr Val Cys

Gly Thr Tyr Ala Phe Gln Pro Lys Cys Thr Tyr Ile Asn Met Leu Thr

Phe Thr Leu Asp Arg Ala Glu Phe Glu Asp Gly Lys Gly Lys Cys Pro

Tyr Asp Pro Ala Lys Gly His Thr Gly Leu Leu Val Asp Gly Glu Leu

Tyr Ser Ala Thr Leu Asn Asn Phe Leu Gly Thr Glu Pro Val Ile Leu

Arg Tyr Met Gly Thr His His Ser Ile Lys Thr Glu Tyr Leu Ala Phe

Trp Leu Asn Glu Pro His Phe Val Gly Ser Ala Phe Val Pro Glu Ser

Val Gly Ser Phe Thr Gly Asp Asp Asp Lys Ile Tyr Phe Phe Phe Ser

Glu Arg Ala Val Glu Tyr Asp Cys Tyr Ser Glu Gln Val Val Ala Arg

Val Ala Arg Val Cys Lys Gly Asp Met Gly Gly Ala Arg Thr Leu Gln

Lys Lys Trp Thr Thr Phe Leu Lys Ala Arg Leu Val Cys Ser Ala Pro

Asp Trp Lys Val Tyr Phe Asn Gln Leu Lys Ala Val His Thr Leu Arg

Gly Ala Ser Trp His Asn Thr Thr Phe Phe Gly Val Phe Gln Ala Arg

Trp Gly Asp Met Asp Leu Ser Ala Val Cys Glu Tyr Gln Leu Glu Gln

Ile Gln Gln Val Phe Glu Gly Pro Tyr Lys Glu Tyr Ser Glu Gln Ala

Gln Lys Trp Ala Arg Tyr Thr Asp Pro Val Pro Ser Pro Arg Pro Gly

Ser Cys Ile Asn Asn Trp His Arg Asp Asn Gly Tyr Thr Ser Ser Leu

Glu Leu Pro Asp Asn Thr Leu Asn Phe Ile Lys Lys His Pro Leu Met

Glu Asp Gln Val Lys Pro Arg Leu Gly Arg Pro Leu Leu Val Lys Lys

Asn Thr Asn Phe Thr His Val Val Ala Asp Arg Val Pro Gly Leu Asp

Gly Ala Thr Tyr Thr Val Leu Phe Ile Gly Thr Gly Asp Gly Trp Leu

Leu Lys Ala Val Ser Leu Gly Pro Trp Ile His Met Val Glu Glu Leu

Gln Val Phe Asp Gln Glu Pro Val Glu Ser Leu Val Leu Ser Gln Ser

Lys Lys Val Leu Phe Ala Gly Ser Arg Ser Gln Leu Val Gln Leu Ser

Leu Ala Asp Cys Thr Lys Tyr Arg Phe Cys Val Asp Cys Val Leu Ala

Arg Asp Pro Tyr Cys Ala Trp Asn Val Asn Thr Ser Arg Cys Val Ala

Thr Thr Ser Gly Arg Ser Gly Ser Phe Leu Val Gln His Val Ala Asn

Leu Asp Thr Ser Lys Met Cys Asn Gln Tyr Gly Ile Lys Lys Val Arg

Ser Ile Pro Lys Asn Ile Thr Val Val Ser Gly Thr Asp Leu Val Leu

Pro Cys His Leu Ser Ser Asn Leu Ala His Ala His Trp Thr Phe Gly

Ser Gln Asp Leu Pro Ala Glu Gln Pro Gly Ser Phe Leu Tyr Asp Thr

Gly Leu Gln Ala Leu Val Val Met Ala Ala Gln Ser Arg His Ser Gly

Pro Tyr Arg Cys Tyr Ser Glu Glu Gln Gly Thr Arg Leu Ala Ala Glu

Ser Tyr Leu Val Ala Val Val Ala Gly Ser Ser Val Thr Leu Glu Ala

Arg Ala Pro Leu Glu Asn Leu Gly Leu Val Trp Leu Ala Val Val Ala

Leu Gly Ala Val Cys Leu Val Leu Leu Leu Leu Val Leu Ser Leu Arg

Arg Arg Leu Arg Glu Glu Leu Glu Lys Gly Ala Lys Ala Ser Glu Arg

Thr Leu Val Tyr Pro Leu Glu Leu Pro Lys Glu Pro Ala Ser Pro Pro

Phe Arg Pro Gly Pro Glu Thr Asp Glu Lys Leu Trp Asp Pro Val Gly

Tyr Tyr Tyr Ser Asp Gly Ser Leu Lys Ile Val Pro Gly His Ala Gly

Gly Ser Gly His Pro Leu Pro Glu Leu Ala Asp Glu Leu Arg Arg Lys

Leu Gln Gln Arg Gln Pro Leu Pro Asp Ser Asn Pro Glu Glu Ser Ser

Val

<210> 24
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucletide ZC16189

<400> 24
gtgaatatcg ggtccacaaa ggg 23

<210> 25
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC16188

<400> 25
tgcagcctgt tgaagttctt gttgg 25

<210> 26
<211> 581
<212> DNA
<213> Artificial Sequence

<220>
<223> Zsmf7 PCR fragment

<400> 26
gtgaatatcg gctccacaaa ggggtcctgt ctggataagc gggactgcga gaactacatc 60
actctcctgg agaggcggag tgaggggctg ctggcctgtg gcaccaacgc ccggcacccc 120
agctgctgga acctggtgaa tggcactgtg gtgccacttg gcgagatgag aggctacgcc 180
cccttcagcc cggacgagaa ctccctggtt ctgtttgaag ggacgaggt gtattccacc 240

atccggaagc aggaatacaa tgggaagatc cctcggttcc gccgcatccg gggcgagagt 300
gagctgtaca ccagtgatac tgtcatgcag aacccacagt tcatcaaagc caccatcgtg 360
caccaagacc aggcttacga tgacaagatc tactacttct tccgagagga caatcctgac 420
aagaatcctg aggctcctct caatgtgtcc cgtgtggccc agttgtgcag gggggaccag 480
ggtggggaaa gttcactgtc agtctccaag tggaacactt ttctgaaagc catgctggta 540
tgcagtgatg ctgccaccaa caagaacttc aacaggctgc a 581

<210> 27
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC14298

<400> 27
gtctggaaag gccatgtagg gcag 24

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide ZC14299

<400> 28
cccaggagag tgatgtagtt ctcg 24

<210> 29
<211> 701
<212> PRT
<213> Mus musculus

<400> 29
Met Ala Phe Arg Ala Ile Cys Val Leu Val Gly Val Phe Ile Cys Ser
1 5 10 15
Ile Cys Val Arg Gly Ser Ser Gln Pro Gln Ala Arg Val Tyr Leu Thr
20 25 30
Phe Asp Glu Leu Arg Glu Thr Lys Thr Ser Glu Tyr Phe Ser Leu Ser
35 40 45
His Gln Gln Leu Asp Tyr Arg Ile Leu Leu Met Asp Glu Asp Gln Asp
50 55 60
Arg Ile Tyr Val Gly Ser Lys Asp His Ile Leu Ser Leu Asn Ile Asn
65 70 75 80

Asn Ile Ser Gln Glu Pro Leu Ser Val Phe Trp Pro Ala Ser Thr Ile
85 90 95

Lys Val Glu Glu Cys Lys Met Ala Gly Lys Asp Pro Thr His Gly Cys
100 105 110

Gly Asn Phe Val Arg Val Ile Gln Thr Phe Asn Arg Thr His Leu Tyr
115 120 125

Val Cys Gly Ser Gly Ala Phe Ser Pro Val Cys Thr Tyr Leu Asn Arg
130 135 140

Gly Arg Arg Ser Glu Asp Gln Val Phe Met Ile Asp Ser Lys Cys Glu
145 150 155 160

Ser Gly Lys Gly Arg Cys Ser Phe Asn Pro Asn Val Asn Thr Val Ser
165 170 175

Val Met Ile Asn Glu Glu Leu Phe Ser Gly Met Tyr Ile Asp Phe Met
180 185 190

Gly Thr Asp Ala Ala Ile Phe Arg Ser Leu Thr Lys Arg Met Gln Leu
195 200 205

Arg Thr Asp Gln His Asn Ser Lys Trp Leu Ser Glu Pro Met Phe Val
210 215 220

Asp Ala His Val Ile Pro Asp Gly Thr Asp Pro Asn Asp Ala Lys Val
225 230 235 240

Tyr Phe Phe Phe Lys Glu Arg Leu Thr Asp Asn Asn Arg Ser Thr Lys
245 250 255

Gln Ile His Ser Met Ile Ala Arg Ile Cys Pro Asn Asp Thr Gly Gly
260 265 270

Gln Arg Ser Leu Val Asn Lys Trp Thr Thr Phe Leu Lys Ala Arg Leu
275 280 285

Val Cys Ser Val Thr Asp Glu Asp Gly Pro Glu Thr His Phe Asp Glu
290 295 300

Leu Glu Asp Val Phe Leu Leu Glu Thr Asp Asn Pro Arg Thr Thr Leu
305 310 315 320

Val Tyr Gly Ile Phe Thr Thr Ser Ser Ser Val Phe Lys Gly Ser Ala
325 330 335

Val Cys Val Tyr His Leu Ser Asp Ile Gln Thr Val Phe Asn Gly Pro
340 345 350

Phe Ala His Lys Glu Gly Pro Asn His Gln Leu Ile Ser Tyr Gln Gly
355 360 365

Arg Ile Pro Tyr Pro Arg Pro Gly Thr Cys Pro Gly Gly Ala Phe Thr
370 375 380

Pro Asn Met Arg Thr Thr Lys Asp Phe Pro Asp Asp Val Val Thr Phe
385 390 395 400

Ile Arg Asn His Pro Leu Met Tyr Asn Ser Ile Ser Pro Ile His Arg
405 410 415

Arg Pro Leu Ile Val Arg Ile Gly Thr Asp Tyr Lys Tyr Thr Lys Ile
420 425 430

Ala Val Asp Arg Val Asn Ala Ala Asp Gly Arg Tyr His Val Leu Phe

435 440 445
Leu Gly Thr Asp Arg Gly Thr Val Gln Lys Val Val Val Leu Pro Thr
450 455 460
Asn Ser Ser Ala Ser Gly Glu Leu Ile Leu Glu Glu Leu Glu Val Phe
465 470 475 480
Lys Asn His Val Asp Gly His Ser Cys Ser Arg Phe Tyr Pro Thr Gly
485 490 495
Lys Arg Arg Ser Arg Arg Gln Asp Val Arg His Gly Asn Pro Leu Thr
500 505 510
Gln Cys Arg Gly Phe Asn Leu Lys Ala Tyr Arg Asn Ala Ala Glu Ile
515 520 525
Val Gln Tyr Gly Val Arg Asn Asn Ser Thr Phe Leu Glu Cys Ala Pro
530 535 540
Lys Ser Pro Gln Ala Ser Ile Lys Trp Leu Leu Gln Lys Asp Lys Asp
545 550 555 560
Arg Arg Lys Glu Gly Lys Leu Asn Glu Arg Ile Ile Ala Thr Ser Gln
565 570 575
Gly Leu Leu Ile Arg Ser Val Gln Asp Ser Asp Gln Gly Leu Tyr His
580 585 590
Cys Ile Ala Thr Glu Asn Ser Phe Lys Gln Thr Ile Ala Lys Ile Asn
595 600 605
Phe Lys Val Leu Asp Ser Glu Met Val Ala Val Val Thr Asp Lys Trp
610 615 620
Ser Pro Trp Thr Trp Ala Gly Ser Val Arg Ala Leu Pro Phe His Pro
625 630 635 640
Lys Asp Ile Leu Gly Ala Phe Ser His Ser Glu Met Gln Leu Ile Asn
645 650 655
Gln Tyr Cys Lys Asp Thr Arg Gln Gln Gln Gln Leu Gly Glu Glu Pro
660 665 670
Gln Lys Met Arg Gly Asp Tyr Gly Lys Leu Lys Ala Leu Ile Asn Ser
675 680 685
Arg Lys Ser Arg Asn Arg Arg Asn Gln Leu Pro Glu Ser
690 695 700

<210> 30
<211> 732
<212> PRT
<213> Mus musculus

<400> 30
Glu Glu Arg Leu Ile Arg Lys Phe Glu Ala Glu Asn Ile Ser Asn Tyr
1 5 10 15
Thr Ala Leu Leu Leu Ser Gln Asp Gly Lys Thr Leu Tyr Val Gly Ala
20 25 30
Arg Glu Ala Leu Phe Ala Leu Asn Ser Asn Leu Ser Phe Leu Pro Gly

```
        35                  40                  45
Gly Glu Tyr Gln Glu Leu Leu Trp Ser Ala Asp Ala Asp Arg Lys Leu
    50                  55                  60
Ala Gln Asp Glu Ala Gly Asn Val Ile Leu Glu Asp Gly Lys Gly His
65                  70                  75                  80
Cys Pro Phe Asp Pro Asn Phe Lys Ser Thr Ala Leu Val Val Asp Gly
                85                  90                  95
Glu Leu Tyr Thr Gly Thr Val Ser Ser Phe Gln Gly Asn Asp Pro Ala
            100                 105                 110
Ile Ser Arg Ser Gln Ser Ser Arg Pro Thr Lys Thr Glu Ser Ser Leu
        115                 120                 125
Asn Trp Leu Gln Asp Pro Ala Phe Val Ala Ser Ala Thr Ser Pro Glu
    130                 135                 140
Ser Leu Gly Ser Pro Ile Gly Asp Asp Asp Lys Ile Tyr Phe Phe Phe
145                 150                 155                 160
Ser Glu Thr Gly Gln Glu Phe Glu Phe Phe Glu Asn Thr Ile Val Ser
                165                 170                 175
Arg Val Ala Arg Val Cys Lys Gly Asp Glu Gly Gly Glu Arg Val Leu
            180                 185                 190
Gln Gln Arg Trp Thr Ser Phe Leu Lys Ala Gln Leu Leu Cys Ser Arg
        195                 200                 205
Pro Asp Asp Gly Phe Pro Phe Asn Val Leu Gln Asp Val Phe Thr Leu
    210                 215                 220
Asn Pro Asn Pro Gln Asp Trp Arg Lys Thr Leu Ser Ile Gly Val Phe
225                 230                 235                 240
Thr Ser Gln Trp His Arg Gly Thr Thr Glu Gly Ser Ala Ile Cys Val
                245                 250                 255
Phe Thr Met Asn Asp Val Gln Lys Ala Phe Asp Gly Leu Tyr Lys Lys
            260                 265                 270
Val Asn Arg Glu Thr Gln Gln Trp Tyr Thr Glu Thr His Gln Val Pro
        275                 280                 285
Thr Pro Arg Pro Gly Ala Cys Ile Thr Asn Ser Ala Arg Glu Arg Lys
    290                 295                 300
Ile Asn Ser Ser Leu Gln Leu Pro Asp Arg Val Leu Asn Phe Leu Lys
305                 310                 315                 320
Asp His Phe Leu Met Asp Gly Gln Val Arg Ser Arg Leu Leu Leu Leu
                325                 330                 335
Gln Pro Arg Ala Arg Tyr Gln Arg Val Ala Val His Arg Val Pro Gly
            340                 345                 350
Leu His Ser Thr Tyr Asp Val Leu Phe Leu Gly Thr Gly Asp Gly Arg
        355                 360                 365
Leu His Lys Ala Val Thr Leu Ser Ser Arg Val His Ile Ile Glu Glu
    370                 375                 380
Leu Gln Ile Phe Pro Gln Gly Gln Pro Val Gln Asn Leu Leu Leu Asp
385                 390                 395                 400
```

Ser His Gly Gly Leu Leu Tyr Ala Ser Ser His Ser Gly Val Val Gln
405 410 415
Val Pro Val Ala Asn Cys Ser Leu Tyr Pro Thr Cys Gly Asp Cys Leu
420 425 430
Leu Ala Arg Asp Pro Tyr Cys Ala Trp Thr Gly Ser Ala Cys Arg Leu
435 440 445
Ala Ser Leu Tyr Gln Pro Asp Leu Ala Ser Arg Pro Trp Thr Gln Asp
450 455 460
Ile Glu Gly Ala Ser Val Lys Glu Leu Cys Lys Asn Ser Ser Tyr Lys
465 470 475 480
Ala Arg Phe Leu Val Pro Gly Lys Pro Cys Lys Gln Val Gln Ile Gln
485 490 495
Pro Asn Thr Val Asn Thr Leu Ala Cys Pro Leu Leu Ser Asn Leu Ala
500 505 510
Thr Arg Leu Trp Val His Asn Gly Ala Pro Val Asn Ala Ser Ala Ser
515 520 525
Cys Arg Val Leu Pro Thr Gly Asp Leu Leu Leu Val Gly Ser Gln Gln
530 535 540
Gly Leu Gly Val Phe Gln Cys Trp Ser Ile Glu Glu Gly Phe Gln Gln
545 550 555 560
Leu Val Ala Ser Tyr Cys Pro Glu Val Met Glu Glu Gly Val Met Asp
565 570 575
Gln Lys Asn Gln Arg Asp Gly Thr Pro Val Ile Ile Asn Thr Ser Arg
580 585 590
Val Ser Ala Pro Ala Gly Gly Arg Asp Ser Trp Gly Ala Asp Lys Ser
595 600 605
Tyr Trp Asn Glu Phe Leu Val Met Cys Thr Leu Phe Val Phe Ala Met
610 615 620
Val Leu Leu Phe Leu Phe Phe Leu Tyr Arg His Arg Asp Gly Met Lys
625 630 635 640
Leu Phe Leu Lys Gln Gly Glu Cys Ala Ser Val His Pro Lys Thr Arg
645 650 655
Pro Ile Val Leu Pro Pro Glu Thr Arg Pro Leu Asn Gly Val Gly Pro
660 665 670
Pro Ser Thr Pro Leu Asp His Arg Gly Tyr Gln Ala Leu Ser Asp Ser
675 680 685
Ser Pro Gly Pro Arg Val Phe Thr Glu Ser Glu Lys Arg Pro Leu Ser
690 695 700
Ile Gln Asp Ser Phe Val Glu Val Ser Pro Val Cys Pro Arg Pro Arg
705 710 715 720
Val Arg Leu Gly Ser Glu Ile Arg Asp Ser Val Val
725 730

<210> 31
<211> 699

<212> PRT
<213> Alcelaphine herpesvirus

<400> 31
```
Met Ala Tyr Leu Asn Ala Thr Val Ser Lys Pro Val Ile Ser Leu Leu
1               5                   10                  15
Ser Leu Ser Lys Lys Val Leu Lys Phe Glu His Cys Gly Gly Glu Gly
            20                  25                  30
Gln Cys Leu Gly Leu Ile Thr Glu Phe Val Ile His Pro Ala Ala Met
        35                  40                  45
Gly Thr Leu Cys Val Ser Ile Arg Leu Leu Met Ile Leu Ser Ala Ile
    50                  55                  60
Thr Ala Ala Lys Ser Arg Phe Ile Asp Lys Pro Arg Leu Ile Val Asn
65                  70                  75                  80
Leu Thr Asp Gly Phe Gly Gln His Arg Phe Phe Gly Pro Gln Glu Pro
                85                  90                  95
His Thr Val Leu Phe His Glu Pro Gly Ser Ser Ser Val Trp Val Gly
            100                 105                 110
Gly Arg Gly Lys Val Tyr Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala
        115                 120                 125
Ser Val Arg Thr Val Asn Ile Gly Ser Thr Ala His Glu Pro His Thr
    130                 135                 140
Val Leu Phe His Ser Leu Asn Ser Ser Asp Val Tyr Val Gly Gly Asn
145                 150                 155                 160
Asn Thr Ile Tyr Leu Phe Asp Phe Ala His Ser Ser Asn Ala Ser Thr
                165                 170                 175
Ala Leu Ile Asn Ile Thr Ser Thr His Asn Thr His Arg Leu Ser Ser
            180                 185                 190
Thr Cys Glu Asn Phe Ile Thr Leu Leu His Asn Gln Thr Asp Gly Leu
        195                 200                 205
Leu Ala Cys Gly Thr Asn Ser Gln Lys Pro Ser Cys Trp Leu Ile Asn
    210                 215                 220
Asn Leu Thr Thr Gln Phe Leu Gly Pro Lys Leu Gly Leu Ala Pro Phe
225                 230                 235                 240
Ser Pro Ser Ser Gly Asn Leu Val Leu Phe Asp Gln Asn Asp Thr Tyr
                245                 250                 255
Ser Thr Ile Asn Leu Tyr Lys Ser Leu Ser Gly Ser His Lys Phe Arg
            260                 265                 270
Arg Ile Ala Gly Gln Val Glu Leu Tyr Thr Ser Asp Thr Ala Met His
        275                 280                 285
Arg Pro Gln Phe Val Gln Ala Thr Ala Val His Lys Asn Glu Ser Tyr
    290                 295                 300
Asp Asp Lys Ile Tyr Phe Phe Phe Gln Glu Asn Ser His Ser Asp Phe
305                 310                 315                 320
Lys Gln Phe Pro His Thr Val Pro Arg Val Gly Gln Val Cys Ser Ser
```

```
                  325                     330                     335
Asp Gln Gly Gly Glu Ser Ser Leu Ser Val Tyr Lys Trp Thr Thr Phe
              340                     345                     350
Leu Lys Ala Arg Leu Ala Cys Val Asp Tyr Asp Thr Gly Arg Ile Tyr
          355                     360                     365
Asn Glu Leu Gln Asp Ile Phe Ile Trp Gln Ala Pro Glu Asn Ser Trp
      370                     375                     380
Glu Glu Thr Leu Ile Tyr Gly Leu Phe Leu Ser Pro Trp Asn Phe Ser
385                     390                     395                     400
Ala Val Cys Val Phe Thr Val Lys Asp Ile Asp His Val Phe Lys Thr
                  405                     410                     415
Ser Lys Leu Lys Asn Tyr His His Lys Leu Pro Thr Pro Arg Pro Gly
              420                     425                     430
Gln Cys Met Lys Asn His Gln His Val Pro Thr Glu Thr Phe Gln Val
          435                     440                     445
Ala Asp Arg Tyr Pro Glu Val Ala Asp Pro Val Tyr Gln Lys Asn Asn
      450                     455                     460
Ala Met Phe Pro Ile Ile Gln Ser Lys Tyr Ile Tyr Thr Lys Leu Leu
465                     470                     475                     480
Val Tyr Arg Val Glu Tyr Gly Gly Val Phe Trp Ala Thr Ile Phe Tyr
                  485                     490                     495
Leu Thr Thr Ile Lys Gly Thr Ile His Ile Tyr Val Arg Tyr Glu Asp
              500                     505                     510
Ser Asn Ser Thr Thr Ala Leu Asn Ile Leu Glu Ile Asn Pro Phe Gln
          515                     520                     525
Lys Pro Ala Pro Ile Gln Asn Ile Leu Leu Asp Asn Thr Asn Leu Lys
      530                     535                     540
Leu Tyr Val Asn Ser Glu Trp Glu Val Ser Glu Val Pro Leu Asp Leu
545                     550                     555                     560
Cys Ser Val Tyr Gly Asn Asp Cys Phe Ser Cys Phe Met Ser Arg Asp
                  565                     570                     575
Pro Leu Cys Thr Trp Tyr Asn Asn Thr Cys Ser Phe Lys Gln Arg Val
              580                     585                     590
Ser Val Glu Thr Gly Gly Pro Ala Asn Arg Thr Leu Ser Glu Met Cys
          595                     600                     605
Gly Asp His Tyr Ala Pro Thr Val Val Lys His Gln Val Ser Ile Pro
      610                     615                     620
Leu Leu Ser Asn Ser Tyr Leu Ser Cys Pro Ala Val Ser Asn His Ala
625                     630                     635                     640
Asp Tyr Phe Trp Thr Lys Asp Gly Phe Thr Glu Lys Arg Cys His Val
                  645                     650                     655
Lys Thr His Lys Asn Asp Cys Ile Leu Leu Ile Ala Asn Ser Thr Thr
              660                     665                     670
Ala Thr Asn Gly Thr His Val Cys Asn Met Lys Glu Asp Ser Val Thr
          675                     680                     685
```

Val Lys Leu Leu Glu Val Asn Val Thr Leu Met
690 695

<210> 32
<211> 772
<212> PRT
<213> Mus musculus

<400> 32
Met Gly Trp Phe Thr Gly Ile Ala Cys Leu Phe Trp Gly Val Leu Leu
1 5 10 15
Thr Ala Arg Ala Asn Tyr Ala Asn Gly Lys Asn Asn Val Pro Arg Leu
20 25 30
Lys Leu Ser Tyr Lys Glu Met Leu Glu Ser Asn Asn Val Ile Thr Phe
35 40 45
Asn Gly Leu Ala Asn Ser Ser Ser Tyr His Thr Phe Leu Leu Asp Glu
50 55 60
Glu Arg Ser Arg Leu Tyr Val Gly Ala Lys Asp His Ile Phe Ser Phe
65 70 75 80
Asn Leu Val Asn Ile Lys Asp Phe Gln Lys Ile Val Trp Pro Val Ser
85 90 95
Tyr Thr Arg Arg Asp Glu Cys Lys Trp Ala Gly Lys Asp Ile Leu Lys
100 105 110
Glu Cys Ala Asn Phe Ile Lys Val Leu Glu Ala Tyr Asn Gln Thr His
115 120 125
Leu Tyr Ala Cys Gly Thr Gly Ala Phe His Pro Ile Cys Thr Tyr Ile
130 135 140
Glu Val Gly His His Pro Glu Asp Asn Ile Phe Lys Leu Gln Asp Ser
145 150 155 160
His Phe Glu Asn Gly Arg Gly Lys Ser Pro Tyr Asp Pro Lys Leu Leu
165 170 175
Thr Ala Ser Leu Leu Ile Asp Gly Glu Leu Tyr Ser Gly Thr Ala Ala
180 185 190
Asp Phe Met Gly Arg Asp Phe Ala Ile Phe Arg Thr Leu Gly Asp His
195 200 205
His Pro Ile Arg Thr Glu Gln His Asp Ser Arg Trp Leu Asn Asp Pro
210 215 220
Arg Phe Ile Ser Ala His Leu Ile Pro Glu Ser Asp Asn Pro Glu Asp
225 230 235 240
Asp Lys Val Tyr Phe Phe Phe Arg Glu Asn Ala Ile Gly Gly Glu His
245 250 255
Ser Gly Lys Ala Thr His Ala Arg Ile Gly Gln Ile Cys Lys Asn Asp
260 265 270
Phe Gly Gly His Arg Ser Leu Val Asn Lys Trp Thr Thr Phe Leu Lys
275 280 285

Ala Arg Leu Ile Cys Ser Val Pro Gly Pro Asn Gly Ile Asp Thr His
    290   295   300
Phe Asp Glu Leu Gln Asp Val Phe Leu Met Asn Ser Lys Asp Pro Lys
305   310   315   320
Asn Pro Ile Val Tyr Gly Val Phe Thr Thr Ser Ser Asn Ile Phe Lys
                325   330   335
Gly Ser Ala Val Cys Met Tyr Ser Met Ser Asp Val Arg Arg Val Phe
            340   345   350
Leu Gly Pro Tyr Ala His Arg Asp Gly Pro Asn Tyr Gln Trp Val Pro
        355   360   365
Tyr Gln Gly Arg Val Pro Tyr Pro Arg Pro Gly Thr Cys Pro Ser Lys
    370   375   380
Thr Phe Gly Gly Phe Asp Ser Thr Lys Asp Leu Pro Asp Asp Val Ile
385   390   395   400
Thr Phe Gly Arg Ser His Pro Ala Met Tyr Asn Pro Val Phe Pro Ile
                405   410   415
Asn Asn Arg Pro Ile Met Ile Lys Thr Asp Val Asn Tyr Gln Phe Thr
            420   425   430
Gln Ile Val Val Asp Arg Val Asp Ala Glu Asp Gly Gln Tyr Asp Val
        435   440   445
Met Phe Ile Gly Thr Asp Val Gly Thr Val Leu Lys Val Val Ser Val
    450   455   460
Pro Lys Glu Thr Trp His Asp Leu Glu Glu Val Leu Leu Glu Glu Met
465   470   475   480
Thr Val Phe Arg Glu Pro Thr Thr Ile Ser Ala Met Glu Leu Ser Thr
                485   490   495
Lys Gln Gln Gln Leu Tyr Ile Gly Ser Thr Ala Gly Val Ala Gln Leu
            500   505   510
Pro Leu His Arg Cys Asp Ile Tyr Gly Lys Ala Cys Ala Glu Cys Cys
        515   520   525
Leu Ala Arg Asp Pro Tyr Cys Ala Trp Asp Gly Ser Ser Cys Ser Arg
    530   535   540
Tyr Phe Pro Thr Ala Lys Arg Arg Thr Arg Arg Gln Asp Ile Arg Asn
545   550   555   560
Gly Asp Pro Leu Thr His Cys Ser Asp Leu Glu Asp His Asp Asn His
                565   570   575
His Gly Pro Ser Leu Glu Glu Arg Ile Ile Tyr Gly Val Glu Asn Ser
            580   585   590
Ser Thr Phe Leu Glu Cys Ser Pro Lys Ser Gln Arg Ala Leu Val Tyr
        595   600   605
Trp Gln Phe Gln Arg Arg Asn Arg Arg Ser Lys Arg Glu Ile Arg Met
    610   615   620
Gly Asp His Ile Ile Arg Thr Glu Gln Gly Leu Leu Leu Arg Ser Leu
625   630   635   640
Gln Lys Lys Asp Ser Gly Asn Tyr Leu Cys His Ala Val Glu His Gly

```
                645                     650                     655
Phe Met Gln Thr Leu Leu Lys Val Thr Leu Glu Val Ile Asp Thr Glu
            660                     665                     670
His Leu Glu Glu Leu Leu His Lys Asp Asp Asp Gly Asp Gly Ser Lys
        675                     680                     685
Ile Lys Glu Met Ser Ser Ser Met Thr Pro Ser Gln Lys Val Trp Tyr
    690                     695                     700
Arg Asp Phe Met Gln Leu Ile Asn His Pro Asn Leu Asn Thr Met Asp
705                     710                     715                     720
Glu Phe Cys Glu Gln Val Trp Lys Arg Asp Arg Lys Gln Arg Arg Gln
                725                     730                     735
Arg Pro Gly His Ser Gln Gly Ser Ser Asn Lys Trp Lys His Met Gln
            740                     745                     750
Glu Ser Lys Lys Gly Arg Asn Arg Arg Thr His Glu Phe Glu Arg Ala
        755                     760                     765
Pro Arg Ser Val
    770
```

<210> 33
<211> 691
<212> PRT
<213> Mus musculus

<400> 33

```
Val Met Ile Pro Gly Leu Ala Leu Leu Trp Val Ala Gly Leu Gly Asp
 1               5                  10                  15
Thr Ala Pro Asn Leu Pro Arg Leu Arg Leu Ser Phe Gln Glu Leu Gln
            20                  25                  30
Ala Arg His Gly Val Arg Thr Phe Arg Leu Glu Arg Thr Cys Cys Tyr
        35                  40                  45
Glu Ala Leu Leu Val Asp Glu Glu Arg Gly Arg Leu Phe Val Gly Ala
    50                  55                  60
Glu Asn His Val Ala Ser Leu Ser Leu Asp Asn Ile Ser Lys Arg Ala
65                  70                  75                  80
Lys Lys Leu Ala Trp Pro Ala Pro Val Glu Trp Arg Glu Glu Cys Asn
                85                  90                  95
Trp Ala Gly Lys Asp Ile Gly Thr Glu Cys Met Asn Phe Val Arg Leu
            100                 105                 110
Leu His Ala Tyr Asn His Thr His Leu Leu Ala Cys Arg Thr Gly Ala
        115                 120                 125
Phe His Pro Thr Cys Ala Leu Trp Arg Trp Ala Thr Ala Gly Gly Thr
    130                 135                 140
His Ala Ser Thr Gly Pro Glu Lys Leu Glu Asp Gly Lys Gly Lys Thr
145                 150                 155                 160
Pro Tyr Asp Pro Arg His Arg Pro Pro Ser Val Leu Val Gly Glu Glu
```

165 170 175
Leu Tyr Ser Gly Val Thr Ala Asp Leu Met Gly Arg Asp Phe Thr Ile
180 185 190
Phe Arg Ser Leu Gly Gln Asn Pro Ser Leu Arg Thr Glu Pro His Asp
195 200 205
Ser Arg Trp Leu Asn Glu Pro Lys Phe Val Lys Val Phe Trp Ile Pro
210 215 220
Glu Ser Glu Asn Pro Asp Asp Asp Lys Ile Tyr Phe Phe Phe Arg Glu
225 230 235 240
Ser Ala Val Glu Ala Ala Pro Ala Met Gly Arg Met Ser Val Ser Arg
245 250 255
Val Gly Gln Ile Cys Arg Asn Asp Leu Gly Gly Gln Arg Ser Leu Val
260 265 270
Asn Lys Trp Thr Thr Phe Leu Lys Ala Arg Leu Val Cys Ser Val Pro
275 280 285
Gly Val Glu Gly Asp Thr His Phe Asp Gln Leu Gln Asp Val Phe Leu
290 295 300
Leu Ser Ser Arg Asp Arg Gln Thr Pro Leu Leu Tyr Ala Val Phe Ser
305 310 315 320
Thr Ser Ser Gly Val Phe Gln Gly Ser Ala Val Cys Val Tyr Ser Met
325 330 335
Asn Asp Val Arg Arg Ala Phe Leu Gly Pro Leu Pro His Lys Glu Gly
340 345 350
Pro Thr His Gln Trp Val Ser Tyr Gln Gly Arg Val Pro Tyr Pro Arg
355 360 365
Pro Gly Met Cys Pro Ser Lys Thr Phe Gly Thr Phe Ser Ser Thr Lys
370 375 380
Asp Phe Pro Asp Asp Val Ile Gln Phe Gly Arg Asn His Pro Leu Met
385 390 395 400
Tyr Asn Pro Val Leu Pro Met Gly Gly Arg Pro Leu Phe Leu Gln Val
405 410 415
Gly Ala Gly Tyr Thr Phe Thr Gln Ile Ala Ala Asp Arg Val Ala Ala
420 425 430
Ala Asp Gly His Tyr Asp Val Leu Phe Ile Gly Thr Asp Val Gly Thr
435 440 445
Val Leu Lys Val Ile Ser Val Pro Lys Gly Arg Arg Pro Asn Ser Glu
450 455 460
Gly Leu Leu Leu Glu Glu Leu Gln Val Phe Glu Asp Ser Ala Asp Gly
465 470 475 480
Ser Ala Cys Thr Arg Phe Gln Pro Thr Ala Lys Arg Arg Phe Arg Arg
485 490 495
Gln Asp Ile Arg Asn Gly Asp Pro Ser Thr Leu Cys Ser Gly Asp Ser
500 505 510
Ser His Ser Val Leu Leu Glu Lys Lys Val Leu Gly Val Glu Ser Gly
515 520 525

Ser Ala Phe Leu Glu Cys Glu Pro Arg Ser Leu Gln Ala His Val Gln
530 535 540
Trp Thr Phe Gln Gly Ala Gly Glu Ala Ala His Thr Gln Val Leu Ala
545 550 555 560
Glu Glu Arg Val Glu Arg Thr Ala Arg Gly Leu Leu Leu Arg Gly Leu
565 570 575
Arg Arg Gln Asp Ser Gly Val Tyr Leu Cys Val Ala Val Glu Gln Gly
580 585 590
Phe Ser Gln Pro Leu Arg Arg Leu Val Leu His Val Leu Ser Ala Ala
595 600 605
Gln Ala Glu Arg Leu Ala Arg Ala Glu Glu Ala Ala Ala Pro Ala Pro
610 615 620
Pro Gly Pro Lys Leu Trp Tyr Arg Asp Phe Leu Gln Leu Val Glu Pro
625 630 635 640
Gly Gly Gly Gly Gly Ala Asn Ser Leu Arg Met Cys Arg Pro Gln Pro
645 650 655
Gly His His Ser Val Ala Ala Asp Ser Arg Arg Lys Gly Arg Asn Arg
660 665 670
Arg Met His Val Ser Glu Leu Arg Ala Glu Arg Gly Pro Arg Ser Ala
675 680 685
Ala His Trp
690

<210> 34
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Glu-Glu affinity tag

<400> 34
Glu Glu Tyr Met Pro Met Glu
1 5

<210> 35
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide antigen huzsmf7-2

<400> 35
Asn Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu

1 5 10 15
Asn Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys
20 25 30
Gly Thr Asn Ala
35

<210> 36
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide antigen huzsmf7-3

<400> 36
Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro Lys Pro
1 5 10 15
Asp Lys Cys

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 **:**
C12N 15/12, C07K 14/47, A61K 38/17, C07K 16/18, C12Q 1/68

A3

(11) International Publication Number: **WO 99/45114**

(43) International Publication Date: 10 September 1999 (10.09.99)

(21) International Application Number: PCT/US99/04758

(22) International Filing Date: 3 March 1999 (03.03.99)

(30) Priority Data:
60/076,611 3 March 1998 (03.03.98) US

(71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(72) Inventors: HOLLOWAY, James, L.; 835 N.E. 89th Street, Seattle, WA 98115 (US). LOFTON–DAY, Catherine, E.; 23908 35th Avenue West, Brier, WA 98036 (US).

(74) Agent: LINGENFELTER, Susan, E.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 4 November 1999 (04.11.99)

(54) Title: HUMAN SEMAPHORIN ZSMF–7

(57) Abstract

Semaphorin polypeptides, polynucleotides encoding the polypeptides, and related compositions and methods are disclosed. The polypeptides are expressed in neuronal and lymphatic tissues. The polypeptides may be used within methods for detecting receptors that mediate neurite outgrowth, modulate cellular proliferation and/or differentiation, and immune response.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/04758

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 A61K38/17 C07K16/18 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	EP 0 892 047 A (HOECHST MARION ROUSSEL DE GMBH) 20 January 1999 (1999-01-20) 100% identity between SEQ ID NO 3 of EP892047 and SEQ ID 2 of the application --- -/--	1-36

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 September 1999	14/09/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lejeune, R

1

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/US 99/04758

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	XU X ET AL: "Human semaphorin K1 is glycosylphosphatidylinositol-linked and defines a new subfamily of viral-related semaphorins" JOURNAL OF BIOLOGICAL CHEMISTRY., vol. 273, no. 35, 28 August 1998 (1998-08-28), pages 22428-22434, XP002113886 AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD., US ISSN: 0021-9258 abstract section 'experimental constructs' starting at the last paragraph of page 22428 figure 1 the sequence of sema K1 in figure 1 corresponds to AA 33-666 of seq ID NO 2 ---	1-5, 7-14, 16-23, 29-33,36
P,X	LANGE C ET AL: "New eukaryotic semaphorins with close homology to semaphorins of viruses" GENOMICS., vol. 51, 1 August 1998 (1998-08-01), pages 340-350, XP002113887 SAN DIEGO., US ISSN: 0888-7543 99.8% identity in 2603 bp overlap between the sequence in figure 3 (H-Sema-L) and SEQ ID 1. abstract ---	1-36
A	ENSSER A ET AL: "Alcephaline herpesvirus type 1 has a semaphorin-like gene" JOURNAL OF GENERAL VIROLOGY., vol. 76, 1995, pages 1063-1067, XP002113888 SOCIETY FOR GENERAL MICROBIOLOGY, READING., GB ISSN: 0022-1317 cited in the application the whole document -----	1

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int tional Application No
PCT/US 99/04758

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0892047 A	20-01-1999	DE 19729211 A	14-01-1999
		DE 19805371 A	12-08-1999
		AU 7507698 A	21-01-1999
		CA 2237158 A	09-01-1999
		CN 1209436 A	03-03-1999
		CZ 9802149 A	13-01-1999
		HU 9801511 A	28-05-1999
		PL 327385 A	18-01-1999